

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Chimie Biologie**

Arrêté ministériel : 7 août 2006

Présentée par

Florence MARTIN

Thèse dirigée par **Dr. Yves JOUANNEAU**

préparée au sein du **Laboratoire de Chimie et Biologie des Métaux**

dans **l'École Doctorale de Chimie et Sciences du Vivant**

Exploration de la biodiversité bactérienne dans un sol pollué par les hydrocarbures : analyse par marquage isotopique du potentiel métabolique et de la dynamique des communautés impliquées dans la dégradation.

Thèse soutenue publiquement le **13 octobre 2011**

devant le jury composé de :

Pr. Dominique SCHNEIDER

Professeur de l'Université de Grenoble, Président

Pr. Robert DURAN

Professeur de l'Université de Pau, Rapporteur

Dr. Pascal SIMONET

Directeur de recherche CNRS Centrale Lyon, Rapporteur

Dr Fabrice MARTIN-LAURENT

Directeur de recherche INRA DIJON, Examineur

Dr. Yves JOUANNEAU

Directeur de recherche CNRS Grenoble, Examineur

Pr. Gérard BLAKE

Professeur de l'Université de Savoie, Membre invité



A R.Sturacci
Et à mes parents

Remerciements

Avant tout je voudrais remercier les membres du jury, Robert Duran, Pascal Simonet, Dominique Schneider, Fabrice Martin-Laurent et Gérard Blake pour avoir accepté de juger mes travaux et pour la discussion qu'ils ont apporté sur ce projet.

Un projet de thèse ne se réalise jamais seul, c'est pourquoi je souhaiterai ici remercier toutes les personnes qui ont participé à ces trois années d'aventure scientifique et humaine.

Je souhaiterai tout d'abord remercier Yves Jouanneau, mon directeur de thèse, qui a cru en mes capacités pour relever le défi de ce projet. Durant ces trois années de thèse, l'expression de nos tempéraments respectifs a produit bon nombre d'étincelles, mais au final nos discussions scientifiques et sa disponibilité sans faille m'ont permis d'apprendre et de découvrir énormément tant d'un point de vu technique que scientifique. Je le remercie aussi pour m'avoir transmis son goût pour la recherche et pour m'avoir fait profiter de son Expérience dans le domaine. En dehors du labo Yves m'a aussi permis de découvrir les joies de l'aviron et de la planche à voile, des moments très ressourçant.

Au sein du LCBM, je souhaiterai remercier particulièrement Stéphane Torelli, alias « Boss ». Grâce à son enseignement, j'ai pu mettre en application mes cours de première année d'école dans le domaine de la synthèse organique. Je le remercie pour tous ces conseils et ces 6 mois passés à optimiser la synthèse du phénanthrène marqué. Cela nous a valu quelques sueurs froides, surtout quand le produit, estimé à 1 million d'euro, passait la nuit sur la paillasse. J'ai pu aussi profité de ses conseils d'expert chimiste, quand je cherchais un moyen d'extraire les protéines du sol. Sur ce point merci aussi à Stéphane Ménage, alias « Grand Chef », pour ces conseils.

Je voudrai dire un grand merci à Denis LePaslier du Laboratoire de Metagénomique des Procaryotes au Genoscope à Evry. Après une rencontre en congrès, il a rejoint le projet et grâce à son travail et à celui de ses collaborateurs, Agnès Barbance et Eric Pelletier, nous avons pu obtenir un formidable jeu de données : les librairies et les séquences de gènes

d'ARNr 16S. Merci à lui de m'avoir fait partager son expérience dans le domaine de la métagénomique.

Durant ces trois années, j'ai eu l'occasion de travailler avec Fabrice Martin-Laurent et David Bru du laboratoire Microbiologie du Sol et de l'Environnement à l'INRA de Dijon. A leur côté j'ai appris à utiliser la PCR quantitative. Je les remercie tous les deux pour leur accueil dans leur laboratoire, et pour les discussions enrichissantes que nous avons eu sur le projet. Cela m'a permis d'explorer les données selon de nouvelles perspectives et de développer mon esprit critique sur les méthodes moléculaires que nous utilisons.

Je souhaite aussi remercier Roberto Geremia du Laboratoire d'Ecologie Alpine, à l'Université de Grenoble, qui m'a accueillie dans son laboratoire afin de travailler sur les analyses SSCP. Je remercie Armelle Monier pour son aide et ses précieux conseils lors de la mise en œuvre de la technique, ainsi que Julien Roy, pour sa disponibilité et sa patience, lorsqu'il a fallu traiter les résultats avec un logiciel que je ne maîtrisais que trop peu.

Merci aussi à Gérard Blake du Laboratoire Optimisation de la Conception et Ingénierie de l'Environnement à l'Université de Savoie. Grâce à sa collaboration dans le projet, nous avons pu avoir accès aux lits à macrophytes gérés par l'AREA. Mais surtout, merci à lui pour avoir apporté une vision différente au projet et pour les discussions très intéressantes que nous avons eu. Cela m'a permis d'appréhender des aspects auxquels je n'avais pas forcément pensé.

Je remercie aussi Yohann Couté de la plateforme protéomique du CEA, avec qui nous avons travaillé sur la mise au point d'une méthode d'analyse protéomique à partir d'échantillons de sol.

Merci enfin à Yves Vandenbrouck pour ses conseils dans le domaine de l'analyse phylogénétique. Domaine que je ne connaissais pas et qu'il m'a permis d'appréhender avec des outils simples et très didactiques.

Durant ces trois années de thèse, j'ai aussi eu l'immense chance d'encadrer deux étudiants lors de leur stage de Master II. Laure et Fabien ont donné énormément de leur temps et de leur implication dans les projets qui leurs ont été proposés. A leurs côtés j'ai appris beaucoup sur le plan humain.

Laure, mon Padawan, m'a appris à maîtriser mon côté feu follet et mes épisodes de « râlerie ». Nous avons passé 6 mois de manipes dans une très bonne ambiance que je n'oublierai jamais. Nous avons aussi arpenté la montagne ensemble jusqu'à faire un sommet (2853 m) la veille de sa soutenance.

Fabien, mon Apprenti, pour qui j'ai été moins présente pour cause de rédaction, m'a appris à développer ma didactique. Ses idées scientifiques, souvent décalées, nous ont emmenés très loin dans le champ de la réflexion. Je n'oublierai pas non plus ses petites blagues entre deux manipes.

Enfin je voudrai dire un grand merci à toutes les personnes que j'ai rencontré au LCBM durant ces trois années.

Je commencerai par ma ptite Peg qui m'a accueillie avec son franc parlé et son caractère « je vais droit au but » en me proposant de but en blanc de faire du ski puis de traverser le Vercors... à pieds, sac au dos. Avide de défis, j'ai pu découvrir, dans ses traces, les joies de l'escalade, du ski dans la puff, des rando raquettes les soirs de pleine lune. Je me suis vidé l'esprit, les week-ends, grâce à son amour des sommets et de tous ce dont la nature alpine recèle. Son soutien infailible au cours de ses trois ans m'a permis de gagner en confiance et d'avancer sereinement.

J'ai aussi rencontré Carole et Delphine avec qui j'ai beaucoup discuté de nos aventures de labo. Notre rencontre hebdomadaire sur le terrain de badminton était une soupape supplémentaire après nos longues semaines de manipes.

Et puis au cours de la dernière année, sont arrivées au labo, Anne-Noël, puis Séverine. Deux post-docs au grand cœur qui ont partagé avec moi leur expérience de jeune docteur, et qui m'ont surtout soutenue dans les périodes de manipes un peu délicates pour la première ou de rédaction fort périlleuse pour la seconde.

Merci à elles et à Julien, Sylvain, Sylvia, Fiona, Flora, Geoffrey, Raïssa, Olivier, Christel, Vincent, Babeth et tous les autres qui ont contribué à créer cette très bonne ambiance de travail... mais pas que. Cette ambiance qui fait que venir au labo ce n'ai pas que mener un projet passionnant, mais c'est aussi retrouver un univers riche et animé !

Enfin, en dehors du laboratoire, au cours de ces trois années, j'ai eu la chance de faire partie de l'association des jeunes chercheurs en biologie et chimie de Grenoble, ³⁸Globule. J'ai rencontré, grâce à l'association et à l'organisation du forum BIOTechno, des gens formidables auprès desquels j'ai non seulement développé mon expérience du management et de la gestion de projet, mais j'ai aussi passé de très bons moments de détente post-réunion. Je tiens à remercier les membres des équipes ³⁸Globule, BIODOCS Lyon et ASEC et en particulier Emeline, Mélanie, Julien, Anthony, Mathieu, Rachel, Marie-Eve, Elsa et Nicolas.

Pour terminer, je souhaiterai remercier ma famille et mes amis proches pour leur soutien non seulement tout au long de la thèse mais aussi depuis de nombreuses années.

Résumé

Les hydrocarbures aromatiques polycycliques (HAP) sont des composés ubiquitaires issus de la combustion incomplète de matières organiques. Surtout générés par l'exploitation des produits pétroliers, ils sont à l'origine de pollutions de l'environnement, car ce sont des composés toxiques pour les êtres vivants et pour l'homme en particulier. De nombreuses bactéries capables de dégrader les HAP ont été isolées et étudiées, mais celles qui les dégradent *in situ* sont mal connues. En effet, moins de 5% des bactéries du sol sont cultivables en laboratoire. Le premier objectif de cette étude était d'identifier les bactéries qui dégradent les HAP dans le sol par des méthodes moléculaires indépendantes de la culture. A cette fin, une stratégie de marquage isotopique *in situ* a été mise en œuvre. Elle repose sur l'utilisation du phénanthrène, un HAP à trois cycles, dans lequel l'isotope naturel du carbone a été remplacé par le ^{13}C . Cette molécule a été introduite comme traceur dans des microcosmes contenant du sol provenant d'un bassin de rétention des eaux de ruissellement d'autoroute. Les bactéries capables de métaboliser le HAP ont incorporé le ^{13}C et ont ainsi été identifiées par séquençage des gènes d'ARNr 16S, amplifiés à partir de l'ADN marqué extrait du sol. Les résultats montrent que des Betaprotéobactéries peu étudiées à ce jour, appartenant aux genres *Acidovorax*, *Rhodoferax*, *Hydrogenophaga* et *Thiobacillus*, ainsi que des Rhodocyclaceae, étaient les principaux acteurs de la dégradation du phénanthrène. La prépondérance des Betaprotéobactéries a été établie par des mesures de PCR quantitative. Une analyse dynamique de la diversité bactérienne a montré que celle-ci changeait en fonction de la biodisponibilité du phénanthrène. En outre, la diversité d'arène-dioxygénases impliquées dans la dégradation des HAP a été explorée sur le plan phylogénétique et fonctionnel. Nous avons ainsi détecté des séquences nouvelles, pour la plupart apparentées à des dioxygénases de Sphingomonadales et de Burkholderiales. Grâce à la construction et à l'expression d'enzymes hybrides, il a été possible, pour la première fois, d'associer une activité catalytique d'oxydation des HAP à des séquences partielles de gènes, amplifiées à partir de l'ADN du sol.

Les résultats obtenus et les outils mis au point dans cette étude pourront servir à développer des méthodes de diagnostic et de suivi de la biodégradation de polluants, par exemple dans le cadre d'opérations de bioremédiation de sites pollués par les HAP.

Mots clés : diversité bactérienne, marquage isotopique, gènes d'ARNr16S, arène-dioxygénases, analyse phylogénétique et fonctionnelle, qPCR, protéomique, betaprotéobactéries, HAPs, phénanthrène, biodégradation.

Valorisation des travaux de thèse

Publications Scientifiques

Jouanneau Y., Martin F., Krivobok S., and J.C. Willison (2011) Rin-hydroxylating dioxygenases involved in PAH biodegradation: Structure, function and biodiversity. In (A. I. Koukkou, ed.) Microbial bioremediation of non-metals: current research; Chapter 7, pp. 149-175, Caister Academic Press, Norfolk, UK.

Martin F., Torelli S., Le Paslier D., Barbance B., Martin-Laurent F., Bru D., Geremia R., Blake G. and Jouanneau Y. (2011) Dynamic changes in the bacterial community of a PAH-contaminated soil as revealed by stable isotope probing: soumis à publication

Communications orales ou par poster

Communication orale :

Martin F., Jouanneau Y. Identification in situ des espèces bactériennes capables de dégrader les HAP et étude des enzymes impliquées. 2^e RENCONTRES NATIONALES DE LARECHERCHE SUR LES SITES & SOLS POLLUÉS Pollutions locales et diffuses. ADEME. Octobre 2009. Paris

Posters :

Martin F., Geremia R., Blake G., Coute Y., Jouanneau Y., Les microorganismes acteurs de la dépollution dans un écosystème contaminé aux hydrocarbures aromatiques polycycliques. Rencontre Jeunes Chercheurs de la SFBBM. Août 2009. Nancy

Martin F., Geremia R., Blake G., Coute Y., Jouanneau Y. *In situ* identification of bacteria involved in Polycyclic Aromatic Hydrocarbon biodegradation and proteomic analysis of associated enzymes. BioMicroWorld. Décembre 2009. Lisbonne

Martin F., Malagnoux L., Jouanneau Y. Exploring bacterial dioxygenase diversity in PAH polluted soils using a method which targets catalytic domains. Congrès International de l'Institut des Métaux en Biologie de Grenoble. Septembre 2010. Antrons

Martin F., Torelli S., Le Paslier D., Barbance B., Martin-Laurent F., Bru D., Geremia R., Blake G. and Jouanneau Y. Investigation of bacterial diversity in soil subjected to chronical PAH pollution: towards the design of molecular tools for the management of polluted soils. Rencontres Intersol 2011. Mars 2011. Lyon

Martin F., Le Paslier D., Barbance B., Martin-Laurent F., Bru D., Geremia R., Blake G. and Jouanneau Y. Dynamic response of a soil bacterial community exposed to PAH contamination as monitored by stable isotope probing. Congrès FEMS2011. Juin 2011. Genève

Abréviation

ADNg : Acide Desoxyribo Nucléique
génomique

ADRA : Amplified Ribosomal DNA
Restriction Analysis

APS : Ammonium Persulfate

ARN : Acide Ribo Nucléique

ARNr 16S : sous-unité 16S de l'ARN
ribosomal

BLAST : Basical Local Alignment Search
Tool

CAPS : N-Cyclohexyl-3-
aminopropanesulfonic acid

CE-SSCP : Capillary Electrophoresis
Single Strand Conformation Polymorphism

CHAPS : 3-[(3-Cholamidopropyl)
dimethylammonio]-1-propanesulfonate

CTAB : Cetyltrimethylammonium
bromide

CsCl : Chlorure de Césium

CsTFA : Trifluoroacétate de Césium

Da/kDa : Dalton/kilo Dalton

DGGE : Denaturing Gradient Gel
Electrophoresis

DMSO : Diméthyl Sulfoxyde

dNTP : Désoxyribonucléotide triphosphate

DTT : Dithiothreitol

EDTA : acide éthylène diamine tetra
acétique

FAD : Flavine Adénine Dinucléotide

FISH : Fluorescence *In Situ* hybridization

GC-MS : Gas Chromatography-Mass
Spectrometry

HAP : Hydrocarbure Aromatique
Polycyclique

IMAC : Immobilized Metal Affinity
Chromatography

IPTG : Isopropyl β -D-1-
thiogalactopyranoside

LB : Luria-Bertani Broth

LC-MS/MS : Liquid Chromatography-
Mass Spectrometry tandem

MES : 2-(N-Morpholino) ethanesulfonic
Acid

MOPS : acide 3-(N-
morpholino)propanesulfonique

m/z : ratio masse sur charge d'une
molécule

NAD/NADH : Nicotinamide Adénine
Dinucléotide (oxydé/réduit)

NDO : naphtalène-dioxygénase

NMDS : Nonmetric Multidimensional
Scaling

OTU : Unité Opérationnelle Taxonomique

PCR : Polymerase Chain Reaction

ppm : partie par million ici mg/kg de sol
sec

qPCR : quantitative Polymerase Chain
Reaction

RDP : Ribosomal Database Project

RMN : Résonnance Magnétique Nucléaire

RPM : Rotation Par Minute

SDS : Sodium Dodecyl Sulfate

SDS PAGE : Sodium Dodecyl Sulfate
Polyacrylamide Gel Electrophoresis

SIM : Single Ion Monitoring

SIP : Stable Isotope Probing

SyBr Safe : intercallant de l'ADN
fluorescent

TBE : Tris Borate EDTA

TE : Tris-EDTA

TEMED : Tetramethylethylenediamine

TGGE : Temperature Gradient Gel
Electrophoresis

T-RFLP : Terminal Restriction Length
Polymorphism

Xgal : 5-bromo-4-chloro-3-indolyl-beta-D-
galactopyranoside

Table des matières

Remerciements.....	I
Résumé.....	V
Valorisation des travaux de thèse	IX
Abréviation	XI
Table des matières.....	XIII
Introduction.....	1
 I. Etude Bibliographique	 5
1. Etudier la diversité bactérienne des sols : de la souche isolée aux communautés bactériennes.....	 7
1.1 Diversité bactérienne et classification des microorganismes	8
1.2 Etude de la diversité bactérienne : approches dépendant de la culture.....	12
1.3 Etude de la diversité bactérienne : approches moléculaires classiques	15
1.3.1 Extraction des acides nucléiques du sol	15
1.3.2 Amplification des gènes d'ARNr 16S par PCR et séquençage	16
1.3.3 Profils d'empreintes moléculaires (fingerprinting)	18
1.3.4 Détection des bactéries par microscopie après marquage fluorescent in situ (FISH)	 22
1.3.5 Quantification par PCR en temps réel	22
1.4 Relier les données phylogénétiques aux fonctions physiologiques.....	23
1.5 Les micropuces à ADN comme technique d'identification des bactéries dans l'environnement.....	 26
1.6 La Métagénomique : vers l'identification taxonomique et fonctionnelle haut débit... 27	
1.7 Conclusions	28
2. Les Hydrocarbures Aromatiques Polycycliques (HAP) et leur biodégradation dans les sols	 29
2.1 Propriétés physico-chimiques des HAP	29
2.2 Provenance et devenir des HAP	29
2.3 Toxicité des HAP.....	31
2.4 Biodégradation des HAP par les microorganismes du sol	33
2.4.1 Facteurs influençant la dégradation des HAP.....	33
2.4.2 Facteurs biologiques favorisant l'accessibilité et la dégradation des HAP	33
2.4.3 Adaptation génétique des bactéries	35
2.5 Voies métaboliques de dégradation des HAP.....	36
2.5.1 Voies de dégradation aérobies	36
2.5.1.1.1 Dégradation du naphtalène par les bactéries	36

2.5.1.2 Dégradation du phénanthrène par les bactéries	37
2.5.1.3 Organisation des opérons cataboliques chez les bactéries.....	40
2.5.1.4 Dégradation des HAP par les champignons	41
2.5.1.5 Dégradation des HAP par les microalgues	42
2.5.2 Voies de dégradation anaérobies chez les bactéries	43
2.6 La diversité microbienne dégradant les HAP dans les sols	44
2.6.1 Les Gammaprotéobactéries	44
2.6.2 Les Alphaprotéobactéries	45
2.6.3 Les Betaprotéobactéries.....	45
2.6.4 Les Actinobactéries	46
2.6.5 Identification d'acteurs de la dégradation des HAP par une approche indépendante de la culture	47
2.6.6 Les acteurs de la dégradation anaérobie des HAP.....	47
3. Les arène-dioxygénases	48
3.1 Organisation structurale.....	48
3.1.1 Les transporteurs d'électrons.....	48
3.1.2 L'oxygénase terminale et son mécanisme catalytique	48
3.2 La dioxygénase de <i>Sphingomonas</i> sp. CHY1.....	52
3.2.2 Eléments structuraux de la sous-unité α	52
3.2.3 Spécificité du substrat.....	53
3.3 Diversité des dioxygénases.....	55
3.3.1 Dioxygénases des bactéries à Gram négatif	55
3.3.1.2 Les Gammaprotéobactéries	57
3.3.1.3 Les Alphaprotéobactéries	57
3.3.1.4 Les Betaprotéobactéries.....	58
3.3.2 Dioxygénases des bactéries à Gram positif	58
II. Présentation du travail de thèse	59
III. Matériel et Méthodes.....	65
1. Sols d'étude, échantillonnages, teneur en HAP	67
1.1 Provenance des échantillons.....	67
1.2 Extraction des HAP	67
2. Mise en place des microcosmes pour les expériences de marquage <i>in situ</i> par un isotope stable	69
2.1 Essais préliminaires	69
2.2 Microcosmes pour la série d'expérience SIP1	69

3. Techniques d'analyses physico chimiques	70
3.1 Dosage des HAP et des dérivés hydroxylés par chromatographie gazeuse et spectrométrie de masse	70
3.1.1 Dosage des HAP	70
3.1.2 Dosage des dérivés hydroxylés	71
3.2 Dosage du CO ₂ par chromatographie gazeuse et spectrométrie de masse	71
3.3 Dosage du phénanthrène en solution dans le DMSO	72
3.4 Analyse LC/MS-MS des protéines du sol	73
3.4.1 Digestion des protéines	73
3.4.2 Analyses en chromatographie liquide couplé à un tandem de spectrométrie de masse (nano-LC-MS/MS)	73
3.4.3 Analyses bioinformatiques	74
4. Techniques de microbiologie	74
4.1 Souches bactériennes utilisées	74
4.2 Milieux de culture	76
4.2.1 E.coli et Pseudomonas	76
4.2.1.1 Milieu LB	76
4.2.1.2 Milieu M9	76
4.2.2 Sphingomonas et Mycobacterium	76
4.2.2.1 Milieu MSM	76
4.2.2.2 Milieu PTYG	77
4.2.2.3 Milieu PTYG 20 solide	77
4.2.3 Acidovorax	77
5. Biologie moléculaire	77
5.1 Plasmides utilisés dans cette étude	77
5.2 Préparation ADN génomique	79
5.2.1 Escherichia coli	79
5.2.2 Sphingomonas	79
5.2.3 Mycobacterium	80
5.2.4 Extraction d'ADN à partir du sol	81
5.3 Séparation d'ADN génomique marqué ¹³ C sur gradient de densité CsCl	81
5.3.1 Préparation des gradients	81
5.3.2 Fractionnement du gradient et détection de l'ADN génomique	82
5.4 Amplification, analyse et purification de l'ADN	83
5.5 Propagation et purification de l'ADN plasmidique	84
5.6 Techniques de clonage	85
5.6.1 Préparation des cellules chimio-compétentes	85

5.6.2 Préparation des cellules électrocompétentes	85
5.6.3 Clonage classique	86
5.6.4 Clonage de produits de PCR bouts francs	87
5.6.5 Clonage sans ligase (Polymerase Incomplete Primer Extension, PIPE)	88
5.7 Vérification des clonages.....	89
5.8 Profils de diversité	90
5.8.1 Polymorphisme de longueur des fragments de restriction (Restriction Fragment Length Polymorphism, RFLP)	90
5.8.2 Polymorphisme de conformation des simples brins d'ADN (Single Strand Conformation Polymorphism, SSCP).....	90
5.9 Construction et séquençage des bibliothèques de gènes d'ARNr 16S	91
5.10 Comparaison des domaines catalytiques de dioxygénases.....	92
6. PCR quantitative en temps réel (qPCR).....	92
6.1 Principe de la méthode	92
6.2 Tests d'inhibition.....	94
6.3 Quantification	94
6.4 Traitement des résultats de qPCR.....	95
7. Méthodes de biochimie	95
7.1 Expression de dioxygénases recombinantes.....	95
7.2 Analyse SDS PAGE	96
7.3 Dosages <i>in vivo</i> de l'activité catalytique des dioxygénases	97
7.4 Purification des protéines recombinantes	98
7.4.1 Expression des dioxygénases hybrides.....	98
7.4.2 Purification des dioxygénases recombinantes	99
7.4.3 Caractéristiques spectrales des protéines surproduites	100
7.4.4 Dosage des protéines	100
8. Extraction et détection des protéines issues des échantillons de sol.....	101
8.1 Extraction des bactéries du sol	101
8.2 Extraction des protéines.....	102
IV. Résultats et Discussions	105
Chapitre I. Mise en place du cadre de l'étude : expérimentations préliminaires	107
1. Synthèse du phénanthrène marqué au carbone ¹³ C.....	108
1.1 Modes opératoires.....	108
1.1.1 Synthèse du naphthalène à partir du benzène	109
1.1.2 Synthèse du phénanthrène à partir du naphthalène	111

1.2 Analyse des produits de synthèse par Résonance Magnétique Nucléaire	113
1.2.1 Principes de RMN et matériel utilisé.....	113
1.2.2 Analyse de l'acide 3-(2-naphtoyl) propionique (2)	115
1.2.3 Spectres RMN du phénanthrène	117
1.3 Chromatographie en phase Gazeuse et Spectrométrie de Masse (GC/MS)	121
2. Dosage des HAP dans le sol	122
3. Etude préliminaire en microcosmes exposés au ¹² C-phénanthrène	124
3.1 Préparation des microcosmes.....	124
3.2 Cinétiques comparées de minéralisation dans les microcosmes	124
3.3 Estimation de la diversité bactérienne dans les sols incubés avec le phénanthrène ..	126
3.3.1 Amplification et clonage des gènes d'ARNr 16S, marqueurs de l'identité bactérienne.....	126
3.3.2 Profils de polymorphisme de restriction (RFLP)	128
4. Conclusions.....	128
Chapitre II. Réponse des populations bactériennes du sol à un épisode de contamination au phénanthrène	
1. Cinétique de la minéralisation du ¹³ C-phénanthrène en microcosmes	129
1.1 Description des microcosmes	129
1.1 Impact d'une dose élevée de phénanthrène sur l'activité métabolique des bactéries du sol.....	130
1.2 Cinétique de transformation du ¹³ C-phénanthrène en ¹³ CO ₂	131
2. Effet du phénanthrène sur la composition globale de la communauté bactérienne du sol	135
1.1 Comparaison des profils de diversité bactérienne dans les microcosmes incubés avec ou sans phénanthrène.....	137
1.1.1 Comparaison des profils SSCP normalisés	138
1.1.2 Influence du temps et de la dose d'exposition au phénanthrène sur la diversité bactérienne suivie par SSCP.....	139
2.2 Effet du phénanthrène sur l'abondance relative des phylums bactériens dominant dans le sol.....	141
3. Discussions et conclusions.....	144
Chapitre III. Caractérisation des communautés bactériennes impliquées dans la dégradation du phénanthrène : changements de diversité en fonction du temps et de la concentration de polluant.....	
1. Isolement de l'ADN génomique marqué ¹³ C par ultracentrifugation isopycnique.....	148
1.1 Séparation des formes marquées et non marquées de l'ADN métagénomique.....	148
1.2 Analyse de la résolution de la séparation des deux formes d'ADN	149

2. Structure et dynamique des communautés bactériennes exposées à une forte dose de phénanthrène	152
2.1 Identification des principaux acteurs de la dégradation du phénanthrène à 5 jours ..	152
2.2 Evolution de la diversité bactérienne marquée au cours des 14 jours d'incubation dans les microcosmes SIP2000	153
2.3 Quantification des phylums bactériens impliqués dans la dégradation du phénanthrène	156
3. Structure et dynamique des communautés bactériennes exposées à une faible dose de phénanthrène (SIP200).....	159
4. Analyse statistique de l'ensemble des données de séquençage	161
4.1 Evaluation de la part des séquences effectivement attribuable aux bactéries impliquées dans la dégradation du phénanthrène.....	162
4.2 Estimation de la diversité bactérienne des microcosmes par l'analyse des librairies de séquences ARNr 16S	164
5. Discussions et conclusions.....	168
5.1 Identification des populations bactériennes impliquées dans la dégradation du phénanthrène.....	168
5.2 Flexibilité de la diversité bactérienne pour répondre à un épisode de contamination	172
5.3 Biais liés aux méthodes moléculaires d'analyse des communautés bactériennes	173
 Chapitre IV. Etude de la diversité des arène-dioxygénases impliquées dans la biodégradation des HAP.....	175
1. Etude la diversité phylogénétique et fonctionnelle des arène-dioxygénases issues des microcosmes par le biais de constructions d'enzymes hybrides.....	176
1.1 Principe de la stratégie envisagée	176
1.2 Difficultés rencontrées et nouvelle stratégie mise en œuvre	177
2. Analyse phylogénétique de la diversité des gènes de dioxygénases détectés dans l'ADN marqué.....	179
2.1 Conception des amorces ciblant les gènes de HAP-dioxygénases de type Protéobactéries.....	179
2.2 Diversité des arène-dioxygénases de type Protéobactéries	180
2.3 Diversité des arène-dioxygénases de type Betaprotéobactérie.....	183
3. Construction, surexpression et dosage de dioxygénases hybrides.....	185
3.1 Construction d'un vecteur d'expression pour la dioxygénase modèle de <i>Sphingomonas</i> CHY-1	185
3.2 Difficultés propres à l'approche expérimentale et mises au point techniques.....	187
3.3 Construction d'un nouveau vecteur pour la surexpression de dioxygénases hybrides dans <i>E. coli</i>	187
3.4 Construction des dioxygénases hybrides dans <i>E. coli</i>	189

3.4.1 Clonage de domaines catalytiques de dcstdiox A dans pFMA1 par la méthode PIPE	189
3.4.2 Clonage de séquences dcstdiox A plus courtes dans pFMA1	190
3.5 Surexpression des dioxygénases hybrides dans <i>E. coli</i>	191
3.6 Dosage de l'activité des dioxygénases hybrides et caractérisations préliminaires....	193
3.7 Purification des dioxygénases hybrides.....	194
3.8 Caractérisation de l'activité des domaines catalytiques de Burkholderiales issus des microcosmes	196
4. Comparaison des séquences peptidiques des dioxygénases hybrides, et modélisation de leur poche catalytique	198
5. Discussion et conclusions	201
5.1 Analyse de la diversité phylogénétique des HAP-dioxygénases des Protéobactéries.	201
5.2 Caractérisation fonctionnelle des dioxygénases associées à la dégradation du phénanthrène.....	204
Chapitre V. Approche protéomique ciblant les arène-dioxygénases des bactéries impliquées dans la dégradation des HAP	207
1. Mise au point d'un protocole d'extraction des cellules bactériennes du sol.....	208
1.1 Critères de choix du procédé d'extraction des cellules bactériennes	208
1.2 Optimisation du procédé d'extraction	209
1.3 Choix du protocole le plus adapté au sol étudié	210
2. Mise au point d'un protocole d'isolement des protéines des bactéries du sol	211
3. Identification des protéines extraites du sol par analyse protéomique.....	213
3.1 Identification d'enzymes du métabolisme des HAP de souches pures cultivées sur phénanthrène marqué.....	213
3.2 Détection des protéines issues d'échantillons de sol	214
4. Discussion et conclusions	217
4.1 Extraction et purification des protéines à partir du sol	217
4.2 Détection des protéines marquées par spectrométrie de masse	218
4.3 Détection des protéines de l'extrait de sol	220
V. Conclusions générales et perspectives du travail	221
VI. Références Bibliographiques	227
Liste des annexes	249

Introduction

Les activités industrielles, économiques et sociales génèrent une multitude de composés plus ou moins toxiques pour les organismes vivants. Par exemple, l'extraction, la transformation et l'utilisation des produits pétroliers produisent des composés organiques que l'on retrouve dans le pétrole brut, les fiouls, les essences, les lubrifiants et des composés inorganiques tels que les métaux lourds (arsenic, plomb, chrome, cuivre, nickel, mercure, cadmium). Ces polluants contaminent tous les compartiments de l'environnement : le sol, l'eau, l'air et la biosphère.

Selon la base de données BASOL (<http://basol.ecologie.gouv.fr/>), les hydrocarbures, les hydrocarbures aromatiques polycycliques (HAP), les solvants halogénés et les métaux lourds sont les composés le plus souvent rencontrés, seuls ou en mélanges, sur les sites et sols pollués. Ces composés sont présents au niveau de friches industrielles (comme les anciennes mines ou cokeries), sur des sites toujours en activités (industries pétrolières, chimiques, sidérurgiques), proche des voies de communication routière, et au niveau des sols agricoles notamment. Ces substances plus ou moins mobiles se déplacent soit sous forme particulaire dans l'atmosphère, soit en solution dans les eaux de surface ou souterraines. Elles peuvent alors atteindre différents écosystèmes.

De part leurs caractéristiques physico-chimiques, ces composés présentent des dangers pour l'environnement et les organismes vivants. Ils entrent dans la chaîne alimentaire, et finissent, à termes, par menacer la santé humaine.

Il est donc primordial de mettre en œuvre des moyens de dépollution des sites contaminés. Leur réhabilitation peut être réalisée par traitement physico-chimique sur site ou après excavation des sols. Ces techniques sont coûteuses et non respectueuses des écosystèmes. Cependant, pour certains polluants organiques (les hydrocarbures, les HAP, certains solvants chlorés), une approche par bioremédiation peut être envisagée. Elle utilise le pouvoir épurateur des microorganismes de l'environnement, et présente l'avantage d'être peu invasive et moins onéreuse.

Une autre préoccupation concerne les sites menacés par des contaminations issues du drainage des eaux de ruissellement, tels que les bords de routes ou d'autoroutes par exemple. La loi sur l'eau de 1992, modifiée en 2006, impose aux collectivités territoriales et aux compagnies d'autoroutes de mettre en place des systèmes afin de gérer les eaux pluviales et de ruissellement, de lutter contre les pollutions, et de protéger et préserver les eaux de surface et les eaux souterraines.

La société d'autoroute AREA a mis en place des systèmes de bassins recueillant les eaux de ruissellement de chaussées d'autoroutes. Depuis leur installation, ces bassins de rétention sont exploités sans suivi régulier de leur capacité à décontaminer les eaux avant leur retour dans l'environnement. Ainsi, les processus de biodégradation mis en jeu dans les bassins n'ont pas encore été étudiés.

Dans le cadre de ce projet, nous nous proposons de caractériser le potentiel bactérien de biodégradation des HAP au sein d'un bassin de rétention recevant les eaux de ruissellement de l'autoroute A43 Nord près de Chambéry (Savoie). Cette étude a pour objectif d'identifier les acteurs spécifiques impliqués dans la dégradation des HAP et d'analyser la diversité des enzymes impliquées dans l'attaque initiale de ces molécules possédant plusieurs cycles aromatiques.

La stratégie mise en œuvre, reposant sur une technique de marquage isotopique *in situ*, sera présentée dans le paragraphe qui suit l'étude bibliographique.

I. Etude Bibliographique

1. Etudier la diversité bactérienne des sols : de la souche isolée aux communautés bactériennes

Afin de bien adapter mon propos au contexte de ce projet, j'ai choisi de limiter l'étude bibliographique présentée dans ce chapitre aux bactéries, à l'exclusion des autres catégories de microorganismes telles que les Archea, les virus, les espèces fongiques ou les algues unicellulaires, qui constituent également la biodiversité microbienne des écosystèmes tels que les sols.

La diversité bactérienne des sols est difficilement quantifiable. Les travaux de ces dix dernières années tendent vers un consensus selon lequel un gramme de sol contiendrait environ 10^9 cellules microbiennes, parmi lesquelles on pourrait distinguer de 10^4 à 10^6 espèces différentes (Curtis *et al.*, 2002; Gans *et al.*, 2005; Janssen, 2006; Sait *et al.*, 2002). La description de la diversité bactérienne repose sur un certain nombre de prérequis et doit tenir compte des conditions environnementales et des incertitudes liées aux méthodes d'analyse.

Tout d'abord, une estimation précise et robuste de la diversité d'une population reposant sur une analyse statistique doit être réalisée sur un échantillon représentatif du sol étudié. Gans et ses collègues montrent, à l'aide d'une distribution logarithmique (népérienne), que la couverture de 80% de la diversité nécessiterait le séquençage de 10^6 clones de gènes d'ARNr 16S (Gans *et al.*, 2005). De plus, l'hétérogénéité spatiale des conditions environnementales induit, dans un même sol, des microhabitats dans lesquels la diversité varie (Torsvik *et al.*, 1996). Par exemple, les paramètres physico-chimiques, les gradients de nutriments ou la présence ou non de plantes modifient drastiquement l'organisation des populations bactériennes aussi bien horizontalement que verticalement (Kirk *et al.*, 2004). Enfin, la pollution des sols est à l'origine de fortes variations de diversité. Par exemple, selon le taux de contamination d'un sol par les métaux lourds, la diversité bactérienne pourrait passer de 10^6 à 10^4 voire 10^3 espèces par gramme de sol, pour la plus forte contamination, alors que la population globale reste à 10^9 cellules par gramme de sol (Gans *et al.*, 2005). L'échantillonnage du sol doit donc être effectué selon des critères qui tiennent compte de la variation de ces paramètres environnementaux afin que l'analyse reflète la diversité réelle.

Les méthodes d'analyse de la diversité microbienne dans l'environnement, et plus spécifiquement de la diversité bactérienne des sols, se heurtent au manque de connaissances, aux difficultés techniques d'analyse, à la complexité et à l'hétérogénéité des sols, etc. Nous verrons, dans un premier temps, les méthodes de classification ainsi que les critiques qui en sont faites dans la littérature. Dans un second temps, un tour d'horizons des principales techniques d'étude des populations bactériennes sera présenté.

1.1 Diversité bactérienne et classification des microorganismes

Au début du XXème siècle, la classification des microorganismes reposait sur des critères morphologiques observés par microscopie ou cultures pures sur boîtes de Petri. A cette époque, les microbiologistes étaient conscients que ces méthodes étaient restrictives et insuffisantes pour distinguer les espèces les unes des autres (Janssen, 2006).

Le développement des méthodes moléculaires, dans les années 1990, a permis de décrire plus précisément les microorganismes et d'établir un nouveau système de classification fondé sur une comparaison de la séquence du gène codant la sous-unité 16S de l'ARN ribosomal (ARNr16S ; Figure I.4). En 1992, Carl Woese et ses collègues présentent la première librairie de séquences de gènes d'ARNr16S (Woese *et al.*, 1990). Ce marqueur biologique de l'identité est présent chez tous les procaryotes. Il a servi de pilier à la classification taxonomique des microorganismes jusqu'à ce jour, et par conséquent à la description des communautés bactériennes de l'environnement. Sur cette base, le *Bergey's Manual of Systematic Bacteriology* (Ludwig and Klenk, 2001) définit les groupes taxonomiques selon 7 niveaux principaux – Domaine, Phylum, Classe, Ordre, Famille, Genre, Espèce – certains de ces niveaux pouvant être déclinés en sous-divisions.

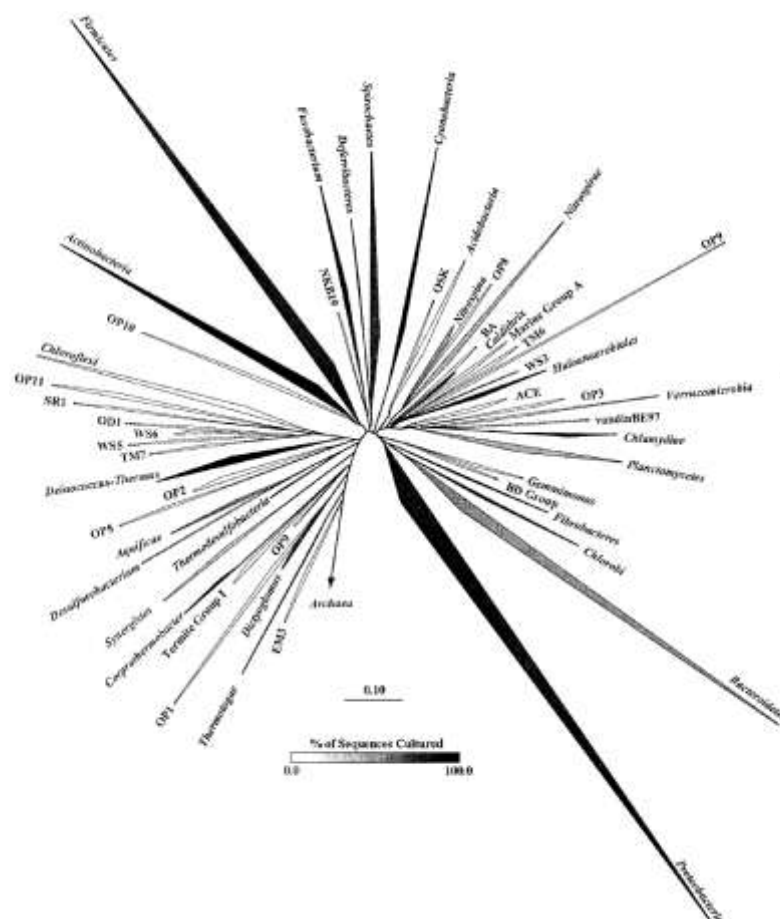


Figure I.1 Arbre phylogénétique des 52 phylums bactériens établis (*italique*) ou candidats (caractères normaux). L'épaisseur des branches reflète la proportion de séquences dans chaque phylum (Schloss and Handelsman, 2004).

L'affiliation de séquences au niveau du phylum est considérée comme fiable et aisée. Actuellement, 52 phylums bactériens majeurs ont été identifiés tous environnements confondus (Figure I.1) (Rappe and Giovannoni, 2003). Dans les sols, certains phylums bactériens sont dominants (Janssen, 2006) et leur abondance relative varie selon les caractéristiques environnementales et les méthodes d'analyse. En comparant 32 librairies de gènes d'ARNr 16S, Janssen (Janssen, 2006) montre qu'en moyenne, 40% des séquences identifiées dans des échantillons de sol appartiennent aux Protéobactéries, 20% aux Acidobactéries et 13% aux Actinobactéries (Figure I.2).

En revanche, l'affiliation de séquences aux niveaux taxonomiques les plus détaillés (genre, espèce) est plus complexe car un grand nombre de genres bactériens n'a pas été caractérisé. Par exemple, dans le phylum des Protéobactéries, bien que les classes Alpha, Beta et Gamma soient bien documentées (528 genres décrits), moins de la moitié des nouvelles séquences qui s'y rapportent peut être affiliée à des genres connus.

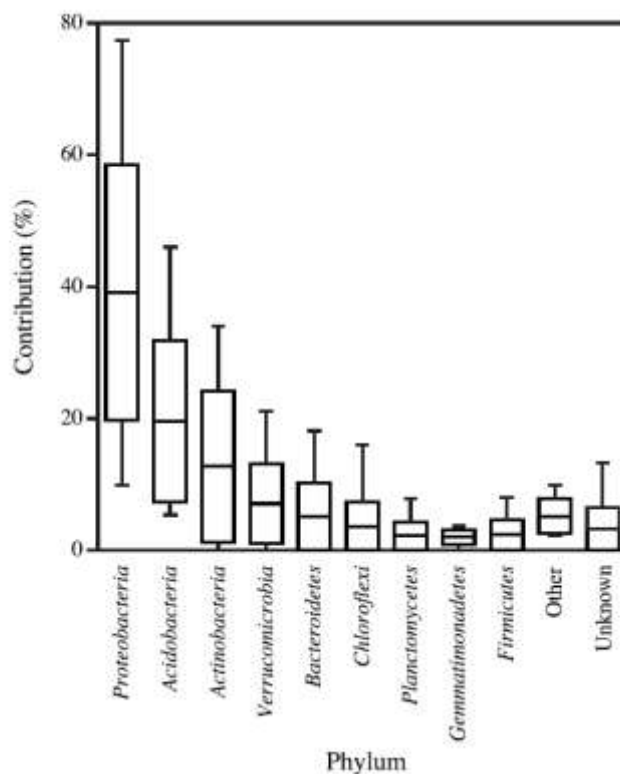


Figure I.2 : Contribution des séquences d'ARNr16S bactériens aux phylums les mieux représentés dans le sol. Les données ont été compilées à partir de 32 librairies de clones d'ARNr16S obtenus par amplification à partir d'ADN du sol. La barre horizontale représente le pourcentage moyen de séquences attribuées au phylum, le rectangle représente l'écart type et les barres verticales représentent les valeurs maximales et minimales (Janssen, 2006)

Malgré ce défaut de connaissance, les séquences des gènes d'ARNr 16S constituent une base robuste et fiable pour évaluer les relations phylogénétiques jusqu'au niveau du genre (Cole *et al.*, 2010). En revanche, la classification au niveau de l'espèce basée sur la seule comparaison des gènes d'ARNr 16S n'est souvent pas pertinente. Ainsi, certaines espèces de *Bacillus*, distinguées sur la base de leur phénotype, présentent des séquences d'ARNr 16S identiques (Maughan and Van der Auwera, 2011).

Dans la classification bactérienne reposant sur le seul critère ARNr 16S, on considère que deux gènes d'ARNr 16S montrant plus de 96% de similarités de séquence appartiennent à des microorganismes de même genre, et que deux gènes ayant plus de 98% de similarité appartiennent à la même espèce. Cependant, en se basant sur d'autres critères comme le taux d'hybridation ADN : ADN, il s'avère que des espèces différentes peuvent avoir des gènes d'ARNr 16S similaires à 99% (Torsvik *et al.*, 1996).

La notion d'espèce étant difficile à définir sur des bases moléculaires, la plupart des auteurs préfèrent donc classer les gènes d'ARNr 16S en unités taxonomiques opérationnelles (OTUs), définies par un seuil de similarité, plutôt que de les affilier à des espèces ou des genres (Schloss and Handelsman, 2005; Torsvik *et al.*, 1996).

Le classement des microorganismes sur la base de leur ARNr 16S ne tient pas compte des variations dans l'organisation qui surviennent par suite d'échanges de gènes entre bactéries. Ainsi, l'adaptation des microorganismes à leur environnement s'accompagne souvent du transfert horizontal de gènes (Juhas *et al.*, 2009).

Bien que les gènes les plus affectés par ce phénomène codent plutôt des enzymes métaboliques, tous les gènes d'un génome peuvent être sujets aux transferts horizontaux. Cependant, ni la fréquence ni le taux des transferts horizontaux des gènes ne sont connus à ce jour. Ainsi, ce phénomène peut générer des taxons divergents portant un gène d'ARNr 16S similaire. Des méthodes de classification alternatives sont actuellement proposées. L'une d'entre elles se base sur la caractérisation des taxons *via* des « gènes de référence » spécifiques dont le nombre varie en fonction inverse du rang taxonomique (Philippot *et al.*, 2010a).

En définitive, les séquences d'ARNr 16S constituent une bonne base d'affiliation taxonomique mais la discrimination en genres et espèces doit s'appuyer sur un séquençage plus systématique d'un panel de « gènes de référence » ou de génomes entiers. De plus, la classification actuelle ne tient compte ni de la nature des écosystèmes d'où proviennent les bactéries, ni des fonctions qu'elles y assurent.

La compréhension de la distribution et de l'organisation des groupes taxonomiques dans différents sols reste un challenge pour la microbiologie environnementale. Certaines méthodes mises en œuvre pour tenter de relever ce défi seront présentées dans les paragraphes suivants.

1.2 Etude de la diversité bactérienne : approches dépendant de la culture

Une des premières techniques d'étude des microorganismes a consisté à inoculer des dilutions d'échantillons naturels dans des milieux de culture artificiels élaborés au laboratoire. Ces milieux, liquides ou solides, requièrent au minimum une source de carbone, une source d'énergie et un accepteur final d'électrons (Hugenholtz, 2002). Les milieux riches sont composés de sources complexes de nutriments – peptone, extrait de levure – métabolisables par un grand nombre de microorganismes, souvent copiotrophes (microorganismes qui ont des demandes élevées en nutriments et qui ont une croissance rapide lorsque les substrats sont abondants (Fierer *et al.*, 2007)).

Les milieux minéraux, contenant un substrat et des sels minéraux, sont souvent utilisés pour isoler par enrichissement des bactéries ayant une fonction métabolique spécifique et conviennent aussi aux microorganismes oligotrophes (microorganismes capables de croître à faible concentration en nutriments grâce à leur haute affinité pour le substrat, leur croissance est plus lente).

Les paramètres de croissance tels que la température, le pH, le temps d'incubation, l'aération, la présence ou non de lumière, varient selon les populations que l'on cherche à détecter. Lorsque des populations spécifiques sont visées, des traitements préalables sont parfois mis en œuvre. La filtration, la centrifugation sur gradient de densité ou la cytométrie en flux (Hugenholtz, 2002) sont des techniques de tri des cellules selon leur taille, leur densité ou d'autres caractéristiques détectables par fluorescence.

Les approches dépendant de la culture sont mises en œuvre pour d'une part, dénombrer les cellules viables (Sait *et al.*, 2002) et d'autre part, isoler par enrichissement des microorganismes doués d'une activité particulière, dégrader un polluant par exemple (Willison, 2004). Si l'identification taxonomique était classiquement basée sur des critères morphologiques et physiologiques, elle est maintenant essentiellement fondée sur l'analyse de gènes de référence (ARNr 16S) ou de génomes (Janssen, 2006). L'isolement de souches bactériennes rend possible l'étude des gènes responsables d'une voie métabolique, et de leur

organisation. Cependant, les approches dépendantes de la culture présentent un certain nombre de biais, comme le fait de favoriser les microorganismes à croissance rapide (Sait *et al.*, 2002). De plus, les cultures de laboratoire reproduisent rarement les conditions qui existent dans l'environnement et qui sont nécessaires à la croissance de la plupart des microorganismes. C'est pourquoi, les bactéries cultivables par des techniques standard ne représentent que 1 à 5% de la diversité observée par des méthodes moléculaires. (Hugenholtz, 2002; Janssen, 2006; Torsvik *et al.*, 1996).

En diversifiant les conditions d'incubation et les milieux de culture, certains auteurs ont réussi à isoler de nouveaux genres bactériens. Par exemple, le remplacement de l'agar par la Gellan Gum a permis de cultiver non plus 5% mais 14% des cellules dénombrées par microscopie. Grâce à cette technique, certaines bactéries apparaissant après 10 semaines de croissance ont été isolées. Leur identification a montré qu'elles étaient affiliées à des microorganismes dont on ne connaissait l'existence que par leur séquence ARNr 16S. De nouveaux taxons d'Acidobactéries, notamment, ont été identifiés (Janssen *et al.*, 2002).

Par ailleurs, un système de culture *in situ* en chambre de diffusion a été mis au point pour permettre d'isoler des microorganismes sur milieu solide, tout en conservant les conditions naturelles d'apport de nutriments (Figure I.3). La majorité des isolats obtenus en chambres de diffusion (70%) n'était pas cultivable sur milieu solide classique (Epstein *et al.*, 2007).

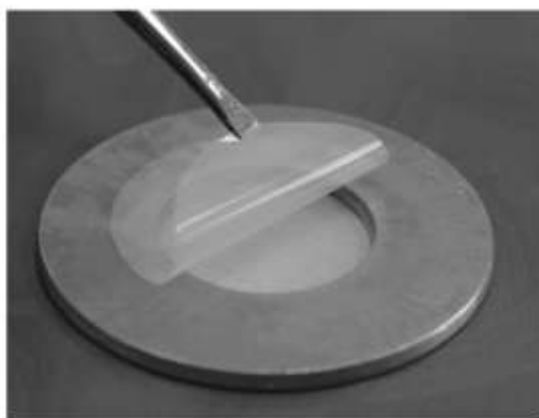


Figure I.3 : Chambre de diffusion développée pour augmenter le potentiel de croissance des microorganismes (Epstein *et al.*, 2007).

Ces exemples montrent que de nouveaux outils de microbiologie sont développés pour isoler et étudier les fonctions de microorganismes méconnus, dont certains pourraient avoir des applications biotechnologiques.

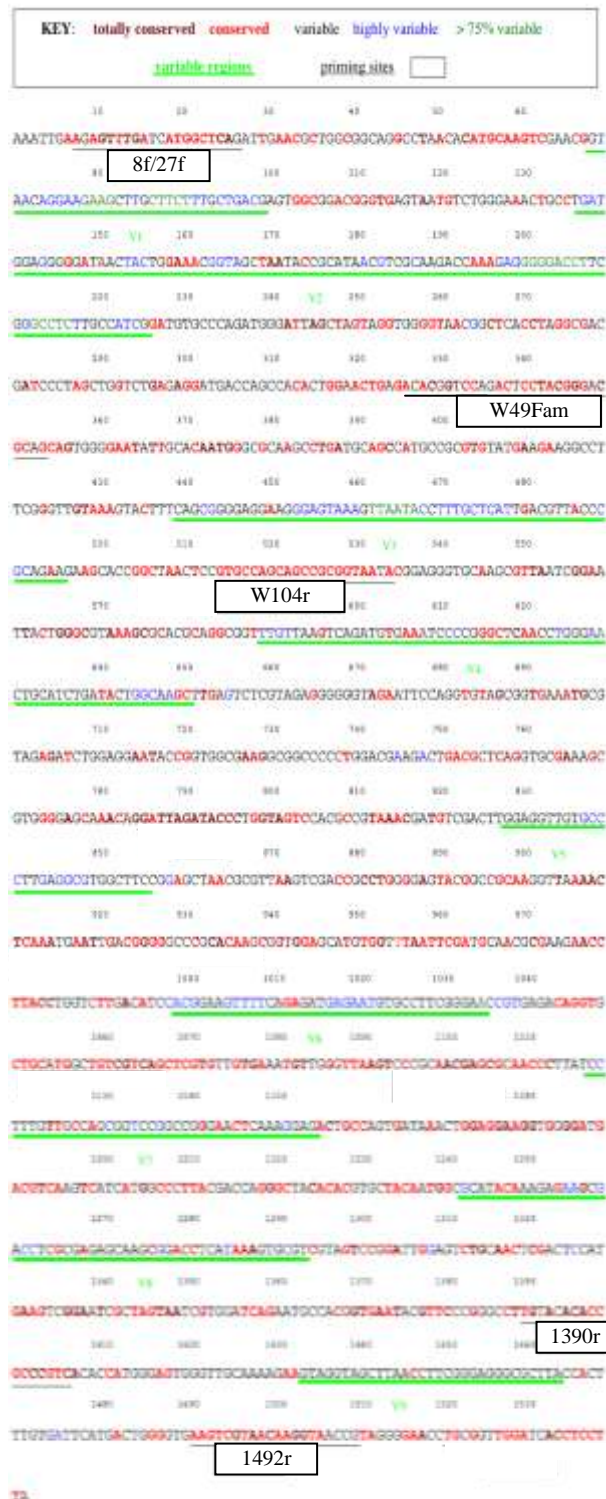


Figure I.4 : Gène de la sous unité 16S de l'ARN ribosomal de *Escherichia coli* les régions variables V1 à V9 utilisées en génotypage. Certaines amorces universelles utilisées lors de cette étude sont représentées. D'après (Baker *et al.*, 2003)

1.3 Etude de la diversité bactérienne : approches moléculaires classiques

1.3.1 Extraction des acides nucléiques du sol

L'extraction des acides nucléiques du sol est une étape cruciale dans l'analyse de la diversité bactérienne puisqu'elle est une source majeure de biais. L'efficacité d'extraction et la pureté de l'ADN ou de l'ARN obtenu sont dépendantes de la composition des sols ou des sédiments d'une part et de la nature des microorganismes d'autre part.

L'extraction d'ADN microbien de sol est problématique du fait de la diversité de sa composition géochimique (argile, sable, matière organique dont acides humiques), de son hétérogénéité (granulométrie, porosité, gradient de nutriments ou de pH, teneur en eau), et de la présence éventuelle de plantes, contribuant à une distribution inégale des microorganismes (Kirk *et al.*, 2004). La stratégie d'échantillonnage, le nombre et le volume des réplicats dépendent donc du sol étudié. Un échantillonnage en gros volumes permet de moyenniser la diversité, alors que la multiplication de petits volumes peut rendre compte de variations dues aux microhabitats du milieu. De plus, les méthodes utilisées pour extraire les acides nucléiques doivent garantir l'élimination de substances inhibitrices des ADN polymérases (Kirk *et al.*, 2004) telles que les matières humiques.

L'extraction des acides nucléiques du sol comporte aussi des biais liés aux bactéries elles-mêmes. Ainsi la résistance mécanique des parois et membranes cellulaires aux techniques de lyse est très variable. Les bactéries à gram positif ont une paroi plus résistante et leur lyse nécessite de mettre en œuvre des traitements drastiques tels que les traitements mécaniques, qui peuvent endommager l'ADN génomique. En revanche, l'utilisation de méthodes trop douces ne donne accès qu'aux bactéries les plus faciles à lyser (Kirk *et al.*, 2004). En outre, certaines bactéries peuvent s'adsorber sur les particules hydrophobes ou s'agréger (Kirk *et al.*, 2004) ce qui diminue leur contribution à l'extrait d'ADN de sol.

Si un seul protocole d'extraction n'est pas suffisant pour capter toute la diversité d'un milieu aussi complexe que le sol, la solution pourrait être d'en utiliser plusieurs comme le préconisent Delmont et collaborateurs. En effet, en comparant deux protocoles d'extraction directe de l'ADN et huit protocoles de lyse de cellules isolées sur gradient de densité, ils

montrent que les populations détectées diffèrent selon les modes d'extraction (Delmont *et al.*, 2011). L'utilisation d'une seule de ces méthodes sous-estime la diversité bactérienne réelle.

1.3.2 Amplification des gènes d'ARNr 16S par PCR et séquençage

Les gènes d'ARNr 16S sont amplifiés à partir de l'ADN génomique extrait des sols soit avec des amorces universelles (exemple Figure I.4), soit avec des amorces plus spécifiques de groupes taxonomiques. Une librairie de clones d'ARNr 16S est ensuite constituée et les gènes sont séquencés par la méthode de Sanger (Sanger *et al.*, 1977). Chaque séquence est alors affiliée à un taxon par comparaison avec les séquences de référence contenues dans les bases de données. Un alignement multiple conduit à la création d'un arbre phylogénétique qui ordonne les séquences expérimentales et calcule des distances relatives par rapport à des séquences de référence.

L'analyse des communautés bactériennes et la diversité observée sont influencées par le choix des amorces d'amplification (Martin-Laurent *et al.*, 2001). La comparaison de la séquence des amorces dites « universelles » avec la partie correspondante des gènes d'ARNr16S des bases de données telles que la RDP (Ribosomal Database Project ; (Cole *et al.*, 2009)) montre qu'elles ne s'hybrident parfaitement qu'à une fraction minoritaire des gènes (Tableau I.1). De ce fait, elles introduisent un biais dans l'amplification PCR et donc dans l'analyse de la biodiversité qui en découle.

Outre les différences de taux d'extraction de l'ADN des bactéries selon leur résistance mécanique, d'autres biais sont liés au nombre de copies de gènes d'ARNr16S par cellule (Janssen, 2006) et à l'amplification préférentielle de certains gènes. Les génomes riches en bases AT par exemple sont plus facilement dénaturés. Lors d'une amplification par PCR, les amorces s'hybrideront plus facilement aux gènes ARNr 16S riches en AT qui seront donc préférentiellement amplifiés (Kirk *et al.*, 2004).

Tableau I.1 : Couverture des séquences de la RDP par des amorces dites universelles.
Les valeurs donnent les pourcentages des séquences qui s'hybrident à 100% avec les amorces.

	8F AGAGTTTGATCCTG GCTCAG	1390R GACGGGCGGTGTGT ACAA	27f* AGAGTTTGATCMTG GCTCAG	1492r* TACGGYTACCTTGT TACGACTT
Phylum et Classes	Taux de couverture des séquences de la RDP (%)			
Bacteria	11,54	23,75	14,01	4,81
Acidobacteria	8,88	19,80	14,43	1,99
Firmicutes	15,93	21,13	17,26	3,86
Bacteroidetes	13,18	21,97	15,51	3,34
Actinobacteria	4,97	17,22	6,26	4,85
Chloroflexi	13,15	27,86	18,11	3,23
Gemmatimonadetes	11,51	21,35	16,04	3,75
Planctomycetes	11,96	31,73	15,52	6,21
Spirochaetes	9,38	43,64	10,88	6,85
Verrucomicrobia	18,74	24,92	21,40	4,00
Cyanobacteria	7,86	28,95	11,89	10,55
Unclassified	11,01	18,54	14,86	3,63
Proteobacteria	10,36	28,64	13,95	6,09
<i>Alpha</i>	9,85	32,35	13,57	6,93
<i>Beta</i>	9,54	22,78	11,72	3,86
<i>Delta</i>	14,30	27,72	18,53	3,17
<i>Epsilon</i>	14,08	1,31	17,11	0,49
<i>Gamma</i>	10,18	31,82	14,39	3,04
<i>Unclassified</i>	9,29	22,00	13,03	3,09

* Ces amorces étant dégénérées, le calcul des % a été fait avec chacune des 2 amorces et les résultats ont été sommés.

Enfin, les bibliothèques de clones doivent avoir une dimension suffisante pour analyser la diversité bactérienne de manière statistique. Un minimum d'une centaine de clones est parfois recommandé (Nelson *et al.*, 2010). Un outil statistique a été mis au point afin d'estimer l'effort de séquençage à partir de données expérimentales, les courbes de raréfaction (Figure I.5). Ces graphiques représentent le nombre de taxons (OTUs) identifiés en fonction du nombre de séquences obtenues. En extrapolant la courbe, il est possible d'estimer, au niveau de l'asymptote, le nombre de séquences nécessaires pour observer 95% de la diversité de l'échantillon. Ainsi, Schloss et Handelsman estiment que 10 000 séquences sont nécessaires pour faire une évaluation pertinente de la richesse de la population bactérienne du sol (Schloss and Handelsman, 2005).

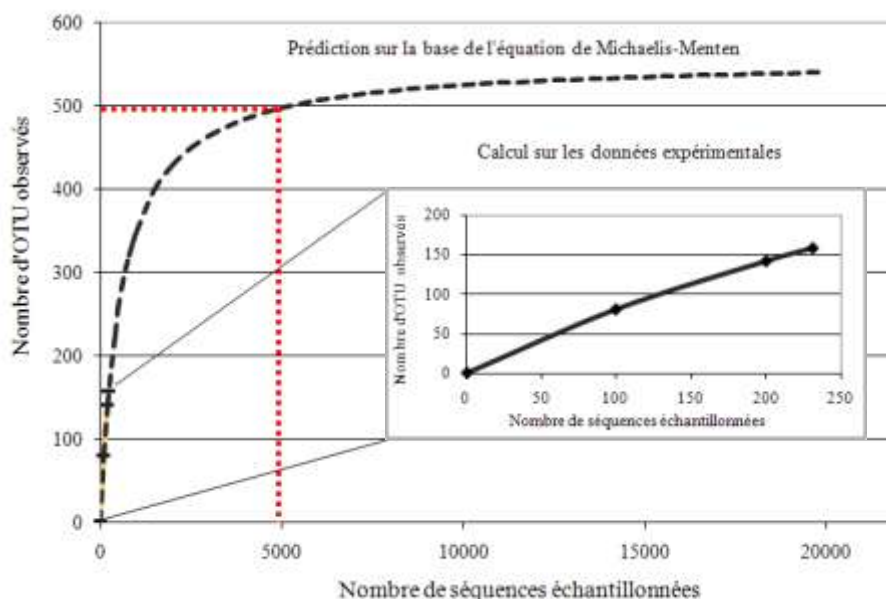


Figure I.5 : Courbe de raréfaction donnant le nombre d'OTUs en fonction du nombre total de séquences d'ARN 16S échantillonnées dans un sol. Le graphique en insert montre la courbe des abondances expérimentales des OTUs en fonction du nombre de séquences obtenues à un niveau de similarité de séquence de 97%. L'extrapolation de cette courbe montre que pour couvrir 90% de la richesse bactérienne du sol, 5000 séquences auraient été nécessaires. Les données proviennent du Chapitre III de la partie Résultats.

Pour toutes les raisons présentées ci-dessus, l'analyse des séquences des gènes d'ARNr 16S conduit à une sous-évaluation de la diversité réelle des échantillons environnementaux. Cependant, l'identification des communautés bactériennes qui sont détectées par cette technique est de plus en plus performante grâce à l'augmentation constante du nombre de séquences contenues dans les bases de données (Nelson *et al.*, 2010).

1.3.3 Profils d'empreintes moléculaires (fingerprinting)

Au cours des 20 dernières années, un large éventail de techniques moléculaires a été développé pour décrire les communautés bactériennes en se basant sur les différences de séquence du gène d'ARN ribosomal (Kirk *et al.*, 2004).

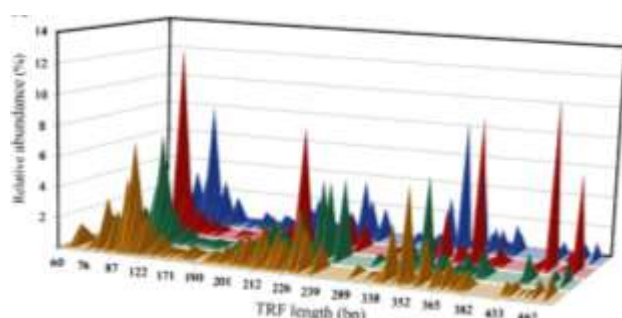
Un premier groupe de techniques compare les profils de fragments obtenus après digestion des gènes d'ARNr 16S par une enzyme à haute fréquence de coupure et séparation par électrophorèse (ARDRA ; Amplified ribosomal DNA restriction analysis) (Bakermans and Madsen, 2002). Le nombre de fragments obtenus, de 4 à 6 par séquence, ne permet d'analyser que les profils de séquences isolées par clonage. Cette méthode peut être utilisée dans le

génotypage d'un groupe de bactéries impliquées dans le métabolisme d'un substrat spécifique (Han *et al.*, 2003). Elle peut aussi permettre de suivre la diversité bactérienne d'un sol utilisé pour la bioremédiation de polluants (Gandolfi *et al.*, 2010).

Une autre méthode utilise le même principe, mais l'analyse se limite à un seul de ces fragments situé à une extrémité. Il est marqué par incorporation d'un oligonucléotide fluorescent lors de l'amplification par PCR. Cette technique connue sous le sigle T-RFLP (Terminal Restriction Fragment Length Polymorphism), appliquée à un mélange de gènes d'ARN 16S, donne un profil représentatif d'une communauté bactérienne (Figure I.6) (Pumphrey and Madsen, 2008a). Contrairement à ARDRA, la technique T-RFLP ne nécessite pas la construction de bibliothèques de clones, qui est un processus qui prend du temps et qui introduit de nombreux biais (PCR, criblage de clones etc.). Ainsi, la méthode T-RFLP permet de suivre l'évolution des populations bactériennes d'un environnement (Lillis *et al.*, 2009) ou de caractériser l'effet d'une pollution sur un milieu naturel (Paisse *et al.*, 2008).

La région inter-génique située entre les séquences codant les sous-unités 23S et 16S peut aussi être employée comme critère distinctif car sa taille varie selon les taxons ((A)RISA ; Automated Ribosomal Intergenic Spacer Analysis) (Martin-Laurent *et al.*, 2001).

Un second groupe de techniques sépare des fragments hypervariables du gène d'ARNr 16S selon les structures secondaires des simples brins obtenus par PCR et dénaturation. Les simples brins sont retenus par paires à l'aide d'une étiquette riche en bases G et C ajoutée lors de la PCR. Après séparation des fragments quasi-simples brins par électrophorèse sur gel de polyacrylamide dénaturant, DGGE (Denaturation Gradient Gel Electrophoresis), ou en présence d'un gradient de température, TGGE (Temperature Gradient Gel Electrophoresis), ces techniques permettent de détecter la présence de populations bactériennes impliquées dans une fonction métabolique donnée (Pumphrey *et al.*, 2011), de suivre des taxons spécifiques lors de la contamination d'environnements naturels (Leys *et al.*, 2004), ou encore d'estimer la diversité des populations bactériennes présentes sur un site pollué de longue date (Cebon *et al.*, 2009).



Profils T-RFLP. Chaque échantillon est représenté par une couleur (Jansson *et al.*, 2007)

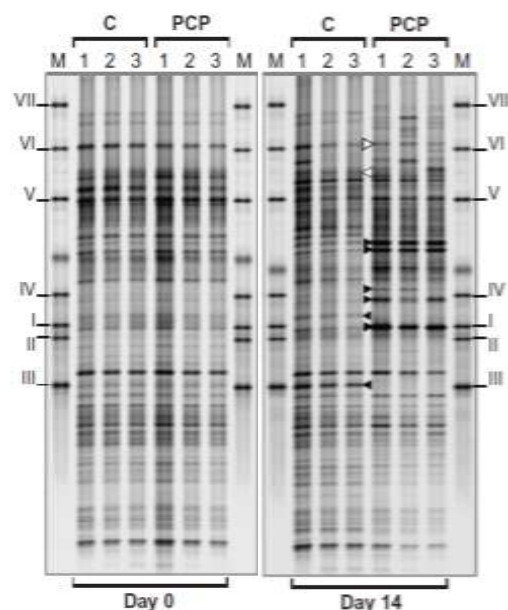


Figure I.6 : Exemples de profils de diversité obtenus par les techniques d'empreintes moléculaires. Le profil T-RFLP a été réalisé dans un appareil qui détecte la fluorescence des fragments terminaux, permettant une automatisation de la technique

Variation de profils DGGE au cours du temps. Les flèches rouges et vertes notent des changements de profils entre 0 et 14 jours (Mahmood *et al.*, 2005)

La méthode CE-SSCP (Capillary Electrophoresis Single Strand Conformational Polymorphism) est une technique basée sur la dénaturation complète de fragments de gènes d'ARNr 16S. Une région variable (Figure I.4) du gène est amplifiée puis dénaturée par traitement thermique. Après refroidissement rapide, les simples brins adoptent des conformations différentes en fonction de leur séquence. Selon la conformation adoptée, ils sont ensuite séparés par électrophorèse capillaire sur une matrice non dénaturante. Une sonde fluorescente, ajoutée à une des extrémités de l'amorce PCR, permet de détecter les brins marqués en sortie de capillaire (Figure I.7) (Zinger *et al.*, 2007). Les profils obtenus permettent de suivre la répartition spatiale de communautés bactériennes (Rodriguez-Blanco *et al.*, 2009) et de comprendre leur structuration en réponse à des conditions environnementales particulières (Zinger *et al.*, 2009).

Les profils de diversité obtenus par ces méthodes permettent d'estimer la diversité, de comparer des communautés bactériennes et de suivre la dynamique des changements spatio-temporels des populations microbiennes des sols (Prosser *et al.*, 2010). Certaines techniques comme la DGGE, la TGGE et la T-RFLP, peuvent en outre fournir des éléments de séquence utilisables pour identifier les taxons majoritaires.

L'automatisation des techniques T-RLFP, ARISA et CE-SSCP augmente les capacités d'analyse, ce qui autorise non seulement la comparaison de plusieurs populations, mais surtout le traitement de plusieurs réplicats, rendant ainsi possible une analyse statistique des données (Ramette, 2007).

Ces techniques présentent néanmoins des limites liées à l'utilisation de la PCR ainsi qu'au manque d'informations précises sur l'identité des microorganismes détectés, même si des séquences partielles sont parfois exploitables. D'autre part, la résolution de ces techniques permet de séparer au mieux une centaine de bandes ou ribotypes, de sorte que chaque bande correspond souvent à plusieurs microorganismes.

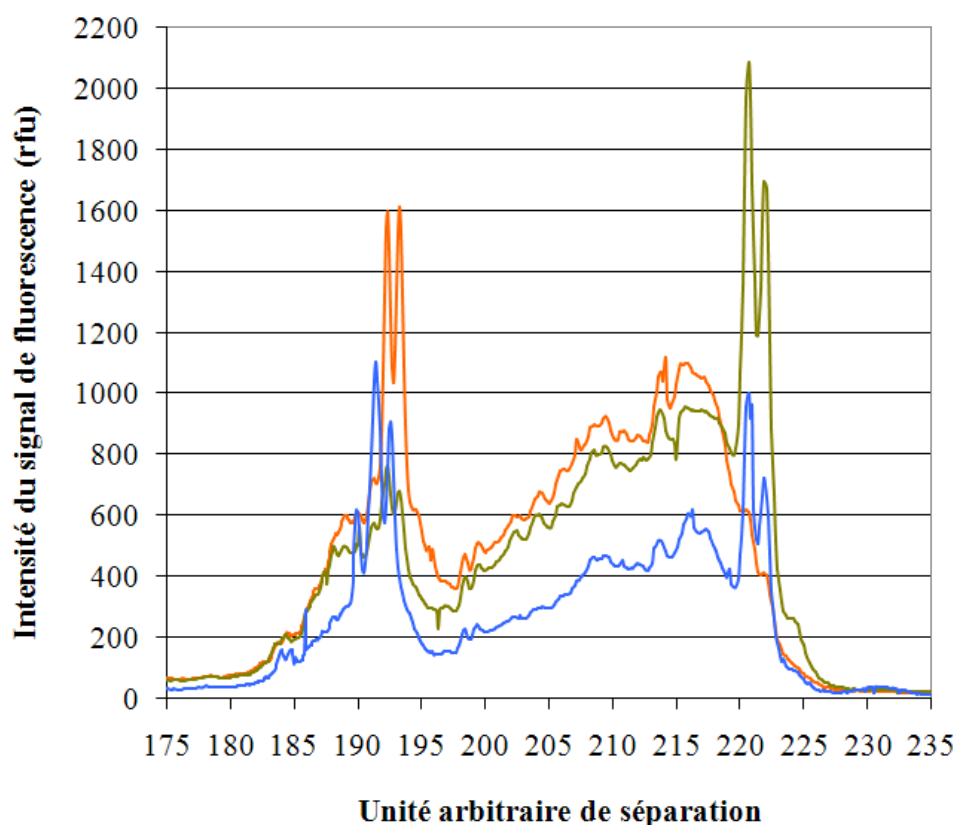


Figure I.7 : Comparaison de profils bactériens d'un sol pollué aux HAP par CE-SSCP. Ces profils représentent la communauté bactérienne d'un sol contaminé par 337 ppm de phénanthrène après 5 jours (orange), 10 jours (vert) et 14 jours.

1.3.4 Détection des bactéries par microscopie après marquage fluorescent *in situ* (FISH)

Les méthodes décrites précédemment font intervenir une étape d'extraction d'ADN génomique, suivie d'une amplification par PCR de marqueurs moléculaires comme préalable aux analyses. Afin de s'affranchir de certains biais qui leur sont associés, des méthodes de surveillance des populations bactériennes par hybridation ont été développées (Amann *et al.*, 1995).

La technique FISH (Fluorescence *in situ* hybridization) consiste à marquer, à l'aide de sondes fluorescentes, des cellules entières afin de les observer dans leur habitat naturel. Les sondes nucléotidiques fluorescentes ciblant des régions variables du gène d'ARN 16S sont introduites dans les bactéries rendues poreuses par certains traitements. Combinée à la microscopie d'épifluorescence, la méthode FISH permet d'étudier les caractères morphologiques des microorganismes et d'analyser leur distribution spatiale. Couplée à la cytométrie de flux, les bactéries appartenant à un taxon peuvent être enrichies par tri sélectif voire même individualisées.

1.3.5 Quantification par PCR en temps réel

La PCR quantitative en temps réel détermine le nombre de copies d'un gène dans un mélange complexe, typiquement un échantillon d'ADN génomique. Le principe de la méthode est détaillé dans le chapitre Matériel et Méthodes. Appliquée à l'écologie microbienne, la technique permet de quantifier des groupes taxonomiques *via* les abondances de leur gène d'ARNr16S spécifique (Fierer *et al.*, 2005). Dans un milieu contaminé, il est ainsi possible de quantifier les communautés bactériennes connues pour leur pouvoir épurateur (Martineau *et al.*, 2010; Singleton *et al.*, 2006), ou de suivre la réaction des populations actives suite à un traitement par biostimulation (Powell *et al.*, 2008). En outre, la qPCR, lorsqu'elle cible une famille de gènes métaboliques ou un gène métabolique spécifique, permet de quantifier le potentiel fonctionnel des microorganismes du sol ou de suivre l'évolution de ce potentiel métabolique en fonction du temps ou des conditions environnementales (Ding *et al.*, 2010; El Azhari *et al.*, 2008; Peng *et al.*, 2010)

Une variante de la technique de qPCR consiste à quantifier les ARN messagers (ARNm), produits de la transcription des gènes. Cette technique désignée par RT-qPCR a le mérite de rendre compte plus fidèlement de l'activité métabolique des bactéries du sol, ou tout au moins de la fonction étudiée. Cependant, elle ne permet pas de quantifier précisément l'activité puisque celle-ci dépend de la part de l'ARNm traduite en protéine active.

Les approches moléculaires décrites ci-dessus reposent sur l'identification de biomarqueurs et sur la mise au point d'amorces pour les amplifier. Les outils développés au cours de ces vingt dernières années ont permis de développer des méthodes de diagnostic et de quantification afin d'obtenir des informations phylogénétiques et fonctionnelles. Il est possible de déterminer l'organisation des microorganismes selon le taux de similarité de séquences ou selon leur potentiel métabolique. En revanche, il est souvent nécessaire d'associer plusieurs biomarqueurs afin de mettre en relation les informations phylogénétiques avec les fonctions métaboliques relatives aux acteurs spécifiques d'un processus biogéochimique donné.

1.4 Relier les données phylogénétiques aux fonctions physiologiques

Dans le but de mieux comprendre le fonctionnement des communautés bactériennes au sein même d'un écosystème, une nouvelle méthode de marquage de la biomasse microbienne par un isotope stable a fait son apparition en 1998 (Stable Isotope Probing, SIP) (Boschker *et al.*, 1998). Deux types de composés marqués ont été majoritairement utilisés : ceux contenant l'isotope ^{13}C du carbone et ceux contenant l'isotope ^{15}N de l'azote. Ici, ne seront présentées que les études impliquant le carbone ^{13}C . A la suite de l'introduction d'un substrat uniformément marqué dans un écosystème microbien, les bactéries capables de l'assimiler vont incorporer l'isotope dans leurs acides nucléiques, acides gras et protéines. L'analyse de marqueurs biologiques enrichis en isotope ^{13}C conduit ensuite à l'identification des bactéries qui ont assimilé le substrat *in situ*.

La première étude SIP a permis d'identifier une population bactérienne impliquée dans la réduction du sulfate *via* l'oxydation de l'acétate (ajouté sous forme ^{13}C) en analysant les profils et la variation des ratios isotopiques des acides gras dérivés des phospho-lipides (PLFA) (Boschker *et al.*, 1998).

En 2000, Radajewski et collaborateurs présentent la première étude basée sur l'analyse de l'ADN marqué avec un isotope stable (ADN-SIP) qui aboutira à l'identification des microorganismes impliqués dans la dégradation du méthanol (Radajewski *et al.*, 2000). Ils utilisent la différence de densité entre l'ADN marqué et l'ADN naturel pour les séparer par ultracentrifugation sur gradient de chlorure de césium. Puis l'amplification par PCR et le séquençage des gènes d'ARNr 16S à partir de l'ADN marqué identifie des microorganismes de la classe des Alphaprotéobactéries ainsi que des Acidobactéries comme les acteurs de la dégradation du méthanol *in situ*.

En 2002, la méthode SIP est employée pour marquer les microorganismes impliqués dans le métabolisme du phénol (Manefield *et al.*, 2002). Après transcription inverse des ARNr marqués (ARN-SIP), DGGE et séquençage du ribotype dominant, l'étude montre que la dégradation du phénol est majoritairement assurée par des bactéries du genre *Thauera*.

La méthode SIP couplée à l'analyse des acides gras membranaires (PLFA-SIP) est d'une grande sensibilité car la détection se fait par chromatographie en phase gazeuse couplée à la spectrométrie de masse, et ne nécessite pas l'extraction et la purification des acides gras marqués (Neufeld *et al.*, 2007). Cependant, le ratio des acides gras n'étant pas un critère suffisamment sélectif, l'identification précise des microorganismes est rarement possible.

L'approche SIP appliquée aux acides nucléiques nécessite l'incorporation de l'isotope stable dans l'ADN lors de sa réplication et nécessite donc la division cellulaire. Les microorganismes en croissance sont les seuls marqués, mais le marquage de l'ADN se fait plus lentement que celui de l'ARN ou des acides gras (Manefield *et al.*, 2002). La méthode ADN-SIP est donc jugée moins sensible que l'approche PLFA-SIP (Neufeld *et al.*, 2007; Prosser *et al.*, 2010). Cependant, l'identification taxonomique des microorganismes ciblés est nettement plus résolutive à partir de l'ADN qu'à partir des acides gras. De plus, cette technique a le potentiel de renseigner sur l'ensemble des capacités métaboliques des bactéries responsables de la fonction étudiée.

Enfin, l'ARN-SIP présente une sensibilité intermédiaire par rapport aux deux autres méthodes (Neufeld *et al.*, 2007). Cette approche nécessite moins d'acides nucléiques pour une bonne séparation des ARN marqués. En effet, l'utilisation de plus d'un microgramme d'ARN par

gradient de CsTFA (trifluoroacétate de césium) provoque une agrégation et une séparation incomplète de l'ARN (Lueders *et al.*, 2004). L'approche SIP appliquée aux ARN ribosomiaux est la plus utilisée des méthodes ARN-SIP. La production de ces ARN est directement liée à la croissance de la population bactérienne qui peut être quantifiée par RT-qPCR (PCR quantitative après transcription inverse). En revanche peu d'études relatent l'utilisation des ARN messagers pour cibler les gènes métaboliques spécifiquement exprimés. Cela peut être dû à des difficultés techniques liées à l'instabilité et à la moindre abondance des ARNm (Neufeld *et al.*, 2007).

La principale difficulté de mise en œuvre de la méthode ADN-SIP réside dans la séparation des acides nucléiques marqués et non marqués avec une résolution suffisante. En effet, suivant (1) la proportion de la communauté microbienne étudiée capable d'incorporer le substrat marqué, et (2) le taux d'incorporation de l'isotope lourd dans la biomasse, la quantité d'ADN marqué sera variable. A cela, il faut ajouter que la différence de densité entre ADN ^{12}C et ^{13}C (0,04 unité) est très proche de celle qui existe naturellement entre un ADN riche en C/G et un ADN riche en A/T (0,03 unité) (Lueders *et al.*, 2004).

Radajewski et collaborateurs ont estimé qu'un enrichissement isotopique de 20 à 50% était nécessaire pour obtenir une résolution suffisante entre l'ADN marqué et le non marqué sur gradient de densité (Radajewski *et al.*, 2003).

Par ailleurs, le métabolisme d'un substrat carboné ne mène pas toujours à sa minéralisation complète. Certaines bactéries produisent des métabolites intermédiaires qu'elles excrètent dans leur environnement. D'autres microorganismes peuvent alors incorporer le carbone ^{13}C provenant de ces métabolites dans leur ADN. Dans ce cas, les populations impliquées dans le métabolisme primaire du substrat ne seront pas les seules à être détectées. De même, certains microorganismes trouvent leurs sources de carbone dans les produits de dégradation relâchés dans l'écosystème à la mort des bactéries qui ont métabolisé le substrat marqué. Le transfert éventuel de carbone marqué à ces populations secondaires, ou « cross feeding », est un phénomène qui prend de l'importance à mesure que le temps d'incubation augmente. Ces phénomènes doivent être pris en compte lors de la mise œuvre des expériences de marquage SIP (DeRito *et al.*, 2005; Neufeld *et al.*, 2007).

1.5 Les micropuces à ADN comme technique d'identification des bactéries dans l'environnement

Les micropuces à ADN ont été développées pour étudier de manière plus commode et plus rapide la diversité microbienne des environnements naturels.

De manière générale, les micropuces à ADN sont composées d'un ensemble de sondes nucléotidiques ciblant des gènes d'intérêt, immobilisées sur des micro-suppôts. L'ADN génomique issu d'échantillons environnementaux est marqué par des sondes fluorescentes, soit directement, soit après amplification (Wang *et al.*, 2011). L'hybridation de fragments d'ADN génomique marqué sur les sondes nucléotidiques immobilisées crée des signaux fluorescents révélateurs de la présence de gènes particuliers dans l'échantillon environnemental.

Plusieurs puces à ADN ont été développées soit pour des analyses phylogénétiques soit pour des analyses fonctionnelles.

Les puces phylogénétiques utilisent des sondes ciblant les gènes d'ARNr 16S. Elles ont été employées pour estimer la diversité bactérienne de sols contaminés (Nemir *et al.*, 2010; Neufeld *et al.*, 2006), et pour caractériser les effets de la pollution du Golfe du Mexique sur la structure des populations bactériennes marines (Hazen *et al.*, 2010).

Les sondes dédiées à l'analyse fonctionnelle sont conçues à partir de gènes spécifiques de processus métaboliques propres à certaines espèces bactériennes.

Par exemple, He et collaborateurs ont développé une micropuce ciblant 150 fonctions métaboliques relatives aux cycles de l'azote, du carbone, du soufre et du phosphore, ainsi que celles associées à la biodégradation de polluants organiques ou à la résistance aux métaux lourds (Wang *et al.*, 2011). A une plus petite échelle, Yergeau et collaborateurs ont analysé le potentiel de biodégradation des hydrocarbures aliphatiques et polyaromatiques de sols de l'Arctique (Yergeau *et al.*, 2009) à l'aide de puces à ADN spécifiques du métabolisme de ces composés. La réalisation d'un diagnostic comparatif entre des sols soumis à différents traitements (*in situ*, *ex situ*, ajout ou pas de nutriments, etc.) a permis de mettre en évidence l'évolution des microorganismes indigènes et d'orienter le choix de bioremédiation le plus adapté à la décontamination du site.

Les micropuces à ADN apparaissent comme des outils performants d'analyse phylogénétique et fonctionnelle de populations bactériennes utilisables pour la surveillance d'écosystèmes, par exemple des sites contaminés.

1.6 La Métagénomique : vers l'identification taxonomique et fonctionnelle haut débit

Le terme de métagénomique a été introduit, dans les années 1990, alors que les microbiologistes s'intéressaient à la caractérisation simultanée du génome de plusieurs microorganismes d'un échantillon environnemental (Hugenholtz and Tyson, 2008). Le développement des techniques associées s'est fait en deux temps. Lors des premières études, l'ADN génomique extrait d'échantillons environnementaux était fragmenté en brins de plusieurs kilobases, puis clonés dans des vecteurs (cosmides, fosmides ou chromosomes artificiels) afin de les séquencer, avant de reconstituer les génomes et des opérons métaboliques. La couverture complète d'un génome nécessitait la construction de plusieurs bibliothèques de fragments de grande taille afin que la séquence issue de l'assemblage soit fiable. Cette approche était très laborieuse et longue à mettre en œuvre (Rajendhran and Gunasekaran, 2008; Shendure and Ji, 2008).

Au début des années 2000, de nouvelles méthodes de séquençage d'ADN haut-débit s'affranchissant des biais d'amplification et de la lourdeur des étapes de clonage ont été mises au point (Ronaghi *et al.*, 1998). Plusieurs revues sur ces techniques ont été publiées (Ahmadian *et al.*, 2006; Medvedev *et al.*, 2009; Metzker, 2010).

Le volume de séquençage produit par ces techniques est considérable, 100 méga bases minimum par lecture, alors que la charge de travail est réduite par rapport aux méthodes classiques de séquençage (Nelson *et al.*, 2010).

Dans ces conditions, l'étape limitante devient le traitement des séquences. Les nombreux outils bioinformatiques, qui continuent à se développer, proposent des algorithmes d'assemblage des séquences en contigs (fragments de génomes) et d'annotation des génomes. Ainsi, certains serveurs, comme la RDP (Ribosomal Database Project (Cole *et al.*, 2009)) proposent des plateformes de traitement en ligne.

La métagénomique haut débit est ainsi un moyen d'analyse très puissant pour découvrir le potentiel métabolique de nouveaux microorganismes, explorer de nouvelles voies métaboliques et identifier de nouveaux biocatalyseurs (Nelson *et al.*, 2010).

1.7 Conclusions

Les méthodes décrites ici ne représentent qu'un panel de ce qui est utilisé pour explorer la diversité taxonomique et fonctionnelle des bactéries du sol.

Il n'existe pas de méthode de choix pour décrire la diversité bactérienne de milieux naturels tels que les sols. Les techniques de profilage renseignent sur la complexité et la réponse dynamique des communautés bactériennes à des facteurs environnementaux, alors que les techniques de séquençage permettent d'accéder à l'identification précise des microorganismes dominants. Les techniques de marquage isotopique permettent d'explorer la diversité de communautés bactériennes au plan fonctionnel, alors que les techniques d'hybridation *in situ* (FISH) ou celles utilisant des micropuces conviennent davantage au monitoring sur site. Enfin, l'avènement des approches indépendantes de la culture pourrait donner un nouvel élan à des approches d'isolement des microorganismes, qui représentent le moyen le plus direct et le plus opérationnel de tirer parti sur le plan biotechnologique de la richesse microbiologique des sols. Chaque méthode apporte un éclairage partiel de la diversité des communautés bactériennes et leur combinaison est susceptible de fournir une image plus réaliste de sa complexité (Bombach *et al.*, 2010; Hill *et al.*, 2000).

2. Les Hydrocarbures Aromatiques Polycycliques (HAP) et leur biodégradation dans les sols

2.1 Propriétés physico-chimiques des HAP

Les hydrocarbures aromatiques polycycliques (HAP) sont composés de deux à six noyaux benzéniques, assemblés de manière linéaire (anthracène), angulaire (phénanthrène) ou condensée (pyrène) (Figure I.8). Suivant le nombre de cycles qu'ils possèdent, on parle d'HAP de faible poids moléculaire pour désigner les composés possédant trois cycles au plus, et d'HAP de haut poids moléculaire pour désigner ceux qui possèdent plus de quatre cycles (Doyle *et al.*, 2008). On regroupe parfois avec les HAP certains composés tricycliques substitués possédant aussi des atomes de soufre, comme le dibenzothiophène, d'azote comme le carbazole, ou d'oxygène comme le dibenzofurane et la dibenzodioxine. L'énergie de résonance créée par le nuage dense d'électrons π stabilise les HAP et les rend résistants aux attaques nucléophiles. Ces molécules sont très hydrophobes, peu volatiles (excepté pour les HAP à deux noyaux comme naphthalène) et ont une faible solubilité dans l'eau (Tableau I.2). Leur hydrophobicité augmente avec le nombre de cycles aromatiques. Ils ont tendance à s'adsorber sur les particules de sol et de sédiments (Johnsen *et al.*, 2005). Leurs propriétés physico-chimiques les rendent récalcitrants à la dégradation naturelle, et donc persistants sur les sites contaminés. A titre d'exemple, la demi-vie du phénanthrène est de 16 à 124 jours, mais celle du benzo[a]pyrène peut atteindre jusqu'à 4 années (Doyle *et al.*, 2008).

2.2 Provenance et devenir des HAP

Les HAP sont des composés issus de la transformation de la matière organique, principalement d'origine végétale. Naturellement émis dans l'atmosphère lors d'éruptions volcaniques, ou de feux de forêts, ils sont aussi contenus dans les combustibles fossiles comme le pétrole ou le charbon. Depuis le début de l'ère industrielle, les HAP émis dans l'environnement sont majoritairement d'origine anthropogénique. Le transport, le chauffage résidentiel, le traitement des déchets par incinération émettent dans l'atmosphère des particules riches en HAP qui se dispersent avant de se déposer sur les sols lors d'épisodes de précipitations. D'autre part, l'exploitation des gisements de pétrole et de charbon, le raffinage, le stockage et les rejets accidentels ou incontrôlés (marées noires) sont sources de

pollution des sols et des milieux marins par les HAP. Les sols des friches industrielles telles que les cokeries, les anciennes usines à gaz ou les usines pétrochimiques sont contaminées par des goudrons riches en HAP. Enfin, certains HAP comme le naphthalène ou le phénanthrène, se retrouvent dans des pesticides, fongicides, détergents ou colorants, et entrent dans l'environnement lors d'activités agricoles ou industrielles (Doyle *et al.*, 2008; Jouanneau *et al.*, 1999). Les HAP sont donc des composés ubiquitaires qui contaminent l'air, les eaux douces ou marines, les sédiments et les sols, à des concentrations variables. Suivant leur exposition, les sols peuvent contenir moins d'un centième de ppm de HAP dans les zones dites 'non contaminées' à plusieurs milliers de ppm sur des sites industriels (Cerniglia, 1992). Les HAP adsorbés aux particules de sol ou associés à la matière organique sont peu mobiles, sauf s'ils sont remis en suspension, et sont difficilement dégradables par les microorganismes.

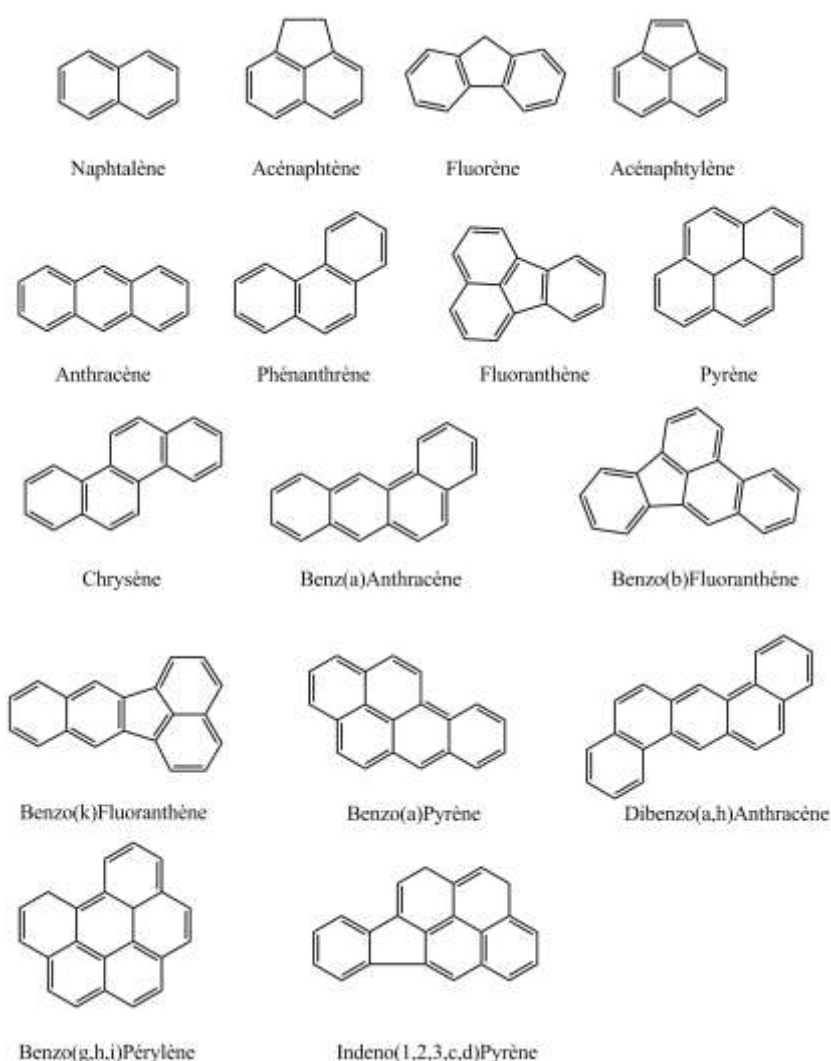


Figure I.8 : Structure des seize HAP prioritaires de la liste de l'EPA (Environmental Protection Agency, USA).

2.3 Toxicité des HAP

En raison de leur caractère liposoluble, les HAP peuvent s'accumuler dans la faune et la flore et tout au long de la chaîne trophique. Chez les mammifères, ils pénètrent dans l'organisme par inhalation, contact cutané ou ingestion, et sont, dans ce cas, absorbés au niveau de la paroi gastro-intestinale. Ils se retrouvent au niveau des tissus adipeux, et de certains organes, notamment le foie, lieu de la détoxification. Des monooxygénases de type cytochromes P450 transforment les HAP par oxydation en époxydes, dérivés phénoliques, quinones, ou en radicaux très réactifs qui forment des adduits avec les protéines et l'ADN.

Les HAP possédant des régions « bay » et des régions « k » (Figure I.9) produisent les époxydes les plus réactifs. Ces sous-produits peuvent se fixer à l'ADN de manière covalente provoquant des mutations, voire des cancers. De même, les HAP peuvent traverser la barrière placentaire et provoquer des dommages chez le fœtus.

Le naphthalène, par exemple, peut provoquer une inhibition de la respiration mitochondriale. Le phénanthrène, qui est le plus petits des HAP possédant une région « bay » et une région « k », est un sensibilisateur de la peau, un allergène, un inducteur faible des échanges de chromatides sœurs et un inhibiteur potentiel de la communication des jonctions cellulaires (Samanta *et al.*, 2002) (Tableau I.2).

L'agence américaine de protection de l'environnement (EPA) et l'agence européenne de sécurité des aliments (EFSA) ont ainsi listé seize HAP comme polluants prioritaires (Figure I.8 ; Tableau I.2).

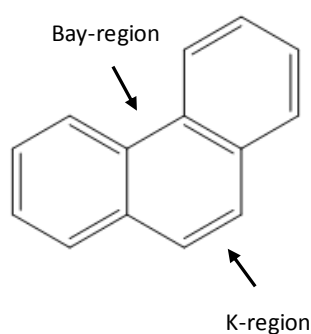


Figure I.9 : Exemple de région « bay » et de région « k » des HAP

Tableau I.2 Propriétés et toxicité des seize HAP prioritaires selon l'EPA.

HAP majoritaires	Formules chimiques	Masse molaire (g/mol)	Solubilité dans l'eau (mg/L)	Toxicité	Concérigène ou mutagène	Commentaires	Sources
Naphtalène	C ₁₀ H ₈	128,16	30	Modérée	Non confirmée	Cancérigène chez les animaux	EPA-TSCA
Acénaphtalène	C ₁₂ H ₁₀	154,21	1,93	Modérée	Constaté		EPA-TSCA
Acénaphtylène	C ₁₂ H ₁₈	154,20	3,93	Modérée	Constaté		EPA-TSCA
Fluorène	C ₁₃ H ₁₀	166,20	1,98	Faible	Constaté	Mutagène pour l'Homme	EPA-TSCA, IARC
Phénanthrène	C ₁₄ H ₁₀	178,20	1,2	Modérée	Constaté	Photosensibilisateur de peau	EPA-TSCA, IARC
Anthracène	C ₁₄ H ₁₀	178,20	0,076	Modérée	Constaté	Candérigène chez les animaux	EPA-TSCA, IARC
Fluoranthène	C ₁₆ H ₁₀	202,26	0,26	Modérée	Constaté	Cancérigène/Retenu par la norme européenne relative à l'eau potable	EPA-TSCA, IARC
Pyrène	C ₁₆ H ₁₀	202,30	0,077	Modérée	Constaté	Domages sur ADN, Mutagène pour l'Homme	EPA-TSCA, IARC
Benz[a]anthracène	C ₁₈ H ₁₂	228,90	0,01	Elevé	Confirmé	Cancérigène, Mutagène pour l'Homme	EPA-TSCA, IARC
Chrysène	C ₁₈ H ₁₂	224,30	2,8 . 10 ⁻³	Elevé	Confirmé	Cancérigène chez les animaux, Mutagène pour l'Homme	EPA-TSCA, IARC
Benzo[b]fluoranthène	C ₂₀ H ₁₂	252,30	0,0012	Elevé	Confirmé	Cancérigène chez les animaux	IARC
Benzo[k]fluoranthène	C ₂₀ H ₁₂	252,30	7,6 . 10 ⁻⁴	Elevé	Confirmé	Cancérigène chez les animaux	IARC
Benzo[a]pyrène	C ₂₀ H ₁₂	252,30	2,3 . 10 ⁻³	Elevé	Confirmé	Cancérigène de l'humain, Mutagène pour l'Homme	EPA-TSCA, IARC
Benzo[g h i]perylène	C ₂₂ H ₁₂	276,34	2,6 . 10 ⁻⁴	Elevé	Constaté		IARC
Dibenzo[a h]anthracène	C ₂₂ H ₁₄	278,35	5 . 10 ⁻⁴	Elevé	Confirmé	Cancérigène chez les animaux, Mutagène pour l'Homme	EPA-TSCA, IARC
Indeno[1,2,3 c d]pyrène	C ₂₂ H ₁₂	276,30	0,062	Elevé	Confirmé	Cancérigène chez les animaux	EPA-TSCA, IARC

Sources : IARC : Centre International de Recherche sur le Cancer

EPA-TSCA : Environmental Protection Agency-Toxic Substances Control Act

Données recensée dans Le rapport de l'INERIS sur les Hydrocarbures Aromatiques Polycycliques dans l'air ambiant, rédigé par Eva LEOZ-GARZIANDIA (Direction des Risques Chroniques) – Décembre 2000

2.4 Biodégradation des HAP par les microorganismes du sol

Certains microorganismes du sol, en particulier des bactéries, sont capables d'obtenir de l'énergie *via* le métabolisme des polluants du sol tels que les HAP (Diaz, 2004). Ils ont développé des stratégies pour utiliser ces molécules comme sources de carbone et d'énergie, assurant la détoxification de l'environnement, ou pour les transformer en substrats métabolisables par d'autres microorganismes (Johnsen *et al.*, 2005). La réduction de la pollution qui en résulte, est exploitée dans des procédés de bioremédiation visant à décontaminer les sites pollués tout en conservant l'intégrité de l'écosystème.

2.4.1 Facteurs influençant la dégradation des HAP

La biodégradation est influencée par des facteurs abiotiques tels que le pH, la température, l'humidité, l'oxygénation, la disponibilité des nutriments, la salinité du sol. La composition du mélange d'HAP détermine leur niveau de solubilité et leur résistance à la dégradation.

De même, l'interaction des polluants avec les constituants du sol détermine l'efficacité de la bio-dépollution. Dans les sols exposés sur de longues périodes, les HAP sont fortement adsorbés aux particules et/ou se logent dans des pores très fins, ce qui les rend inaccessibles aux microorganismes. La biodisponibilité des HAP est donc un facteur essentiel qui bien souvent limite l'efficacité des procédés de bioremédiation (Doyle *et al.*, 2008; Samanta *et al.*, 2002).

2.4.2 Facteurs biologiques favorisant l'accessibilité et la dégradation des HAP

Les bactéries, *via* des voies métaboliques spécifiques, tiennent le rôle central des phénomènes de biodégradation des HAP. Certaines d'entre elles ont développé un système de signalisation, le chimiotactisme, favorisant l'accès aux HAP (Figure I.10) (Samanta *et al.*, 2002). Ces mécanismes interviennent *via* des chimiorécepteurs et des voies de signalisation, et provoquent un déplacement des bactéries selon le gradient de concentration du polluant. Les bactéries s'accumulent alors à l'interface entre le polluant hydrophobe et le milieu hydrophile, provoquant une augmentation du taux de dégradation des composés et souvent une augmentation de leur désorption (Law and Aitken, 2003).

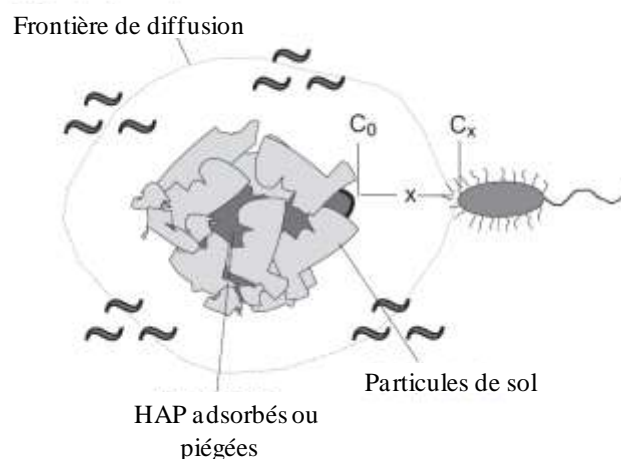


Figure I.10 : Le chimiotactisme permet aux microorganismes de se mouvoir en direction des zones du sol où les polluants sont peu accessibles (Johnsen *et al.*, 2005).

Certains microorganismes sont capables de s'adsorber aux particules riches en polluants grâce à leur paroi hydrophobe (Miyata *et al.*, 2004) et de former des biofilms (Johnsen *et al.*, 2005). D'autres microorganismes produisent des surfactants. Ces molécules ont la propriété d'augmenter la solubilité des composés hydrophobes et d'en améliorer l'accessibilité (Zheng and Obbard, 2002).

Dans le cas de HAP très récalcitrants à la biodégradation comme le benzo[a]pyrène, des processus coopératifs peuvent se mettre en place : les champignons, connus pour leur capacité à dégrader les HAP de haut poids moléculaire (Cerniglia, 1992), amorcent la dégradation. Les bactéries utilisent ensuite les produits de dégradation plus hydrophiles et plus réactifs comme sources de carbone (Boonchan *et al.*, 2000).

De même, l'interaction des bactéries avec les racines des plantes favorise souvent l'activité de dégradation des HAP dans la zone de diffusion des exsudats racinaires ou de l'oxygène. Le co-métabolisme des sources de carbone issues d'extraits de racines permet d'augmenter la capacité de dégradation du benzo[a]pyrène par une souche de *Sphingomonas yanoikuyae* (Rentz *et al.*, 2005). De même, la présence de roseaux a permis de stimuler la dégradation du pyrène dans des sédiments lacustres. Les racines des plantes augmentent, dans ce cas, la diffusion de l'oxygène dans les sédiments et favorisent ainsi la biodégradation bactérienne des HAP (Jouanneau *et al.*, 2005).

Les paramètres environnementaux et le niveau de biodisponibilité des HAP sélectionnent, en quelque sorte, les communautés microbiennes qui seront capables de dégrader les polluants *in situ*. L'influence relative de ces paramètres et leurs relations sont illustrées sur la Figure I.11.

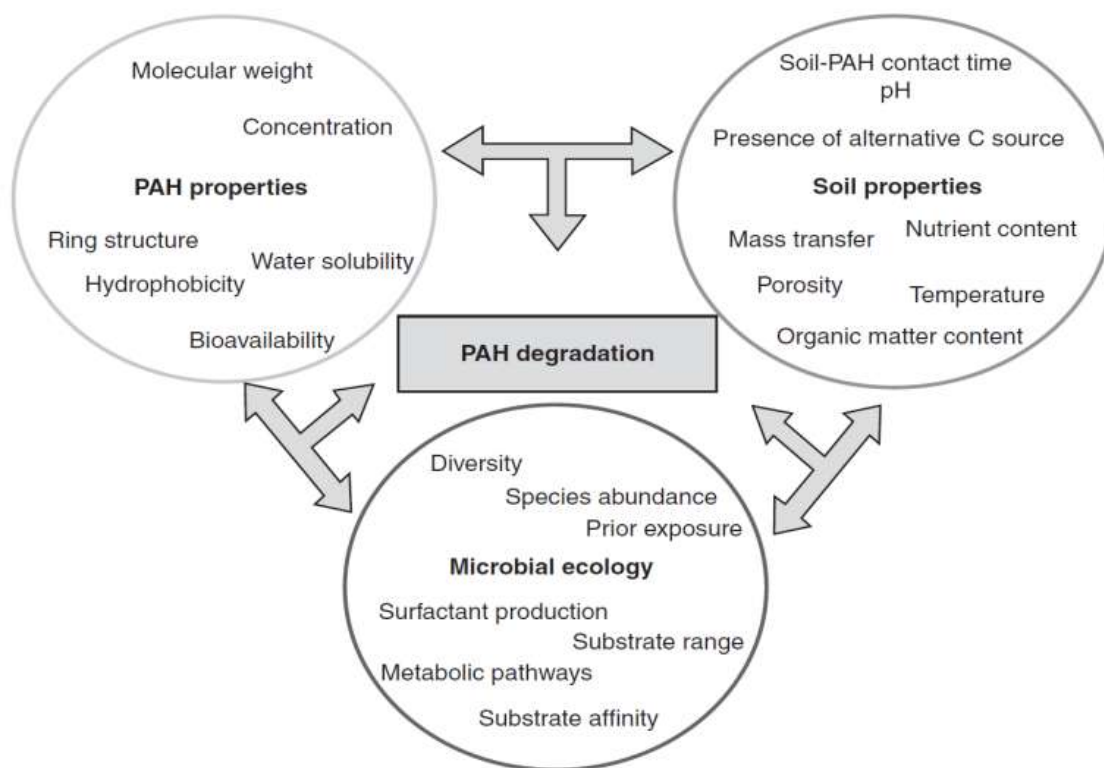


Figure I.11 : Influence des paramètres environnementaux, des propriétés des polluants et de la microflore du sol sur la dégradation des HAP (Doyle *et al.*, 2008).

2.4.3 Adaptation génétique des bactéries

Les microorganismes d'un sol non exposé à des composés contaminants ne possèdent pas forcément la capacité de les métaboliser. Cependant, lorsqu'ils sont exposés à de tels composés, ils sont souvent capables de s'adapter, c'est-à-dire d'acquérir le potentiel métabolique de dégradation de ces polluants par recrutement vertical ou horizontal de gènes spécifiques (George and Hay, 2011).

Le recrutement des voies métaboliques, présentes dans le génome mais non exprimées, peut se faire par des événements de mutation ponctuelle, de réarrangement génétique ou de transposition (recrutement vertical) (Habe and Omori, 2003).

Les microorganismes peuvent aussi acquérir des clusters de gènes cataboliques *via* des éléments mobiles transférés d'un hôte donneur à un hôte receveur (transfert horizontal). Chez les bactéries, les phénomènes de transferts les plus connus sont la conjugaison *via* des plasmides, la transformation ou la transduction (Juhas *et al.*, 2009). Ces phénomènes de transferts horizontaux augmentent les capacités métaboliques des bactéries, notamment en élargissant leur gamme de substrats, et permettent l'adaptation des populations microbiennes à de nouveaux contaminants (van der Meer *et al.*, 1992; van der Meer and Sentchilo, 2003)

2.5 Voies métaboliques de dégradation des HAP

La minéralisation complète des HAP en CO₂ et H₂O suit deux grandes voies cataboliques successives. La première, dite « périphérique » est spécifique du composé ou de la gamme de composés dégradés. Elle produit un métabolite intermédiaire, souvent de type catéchol. Celui-ci entre dans une deuxième voie dite « centrale », commune à de nombreux microorganismes, qui convertit le métabolite en intermédiaires du cycle de Krebs (Diaz, 2004). D'autres voies de dégradation des HAP sont connues chez les champignons et les microalgues, mais celles-ci sont en général incomplètes, de sorte que ces microorganismes ne peuvent utiliser les HAP comme seules sources de carbone. (Doyle *et al.*, 2008; George and Hay, 2011; Peng *et al.*, 2008)

2.5.1 Voies de dégradation aérobies

En conditions aérobies, l'attaque initiale des HAP par les bactéries est réalisée *via* des arène-dioxygénases qui incorporent les deux atomes de la molécule d'oxygène dans l'un des noyaux aromatiques (le mécanisme d'action des dioxygénases sera détaillé dans le paragraphe 3.1.2 de l'Etude bibliographique).

2.5.1.1.1 Dégradation du naphtalène par les bactéries

La voie de dégradation du naphtalène a été étudiée en détail et sert de modèle pour expliquer la dégradation des autres HAP par les bactéries (Figure I.12). La « voie périphérique » débute par l'insertion de deux atomes de l'oxygène moléculaire sur deux carbones adjacents du substrat produisant un *cis*-dihydrodiol. L'attaque initiale est catalysée par une naphtalène-dioxygénase. Le *cis*-dihydrodiol est ensuite ré-aromatisé par une dihydrodiol déshydrogénase,

produisant un intermédiaire dihydroxylé, puis le noyau aromatique est clivé par une extradiol-dioxygénase. Après réarrangement, catalysé par une isomérase, une aldolase conduit à la production de pyruvate et d'acide salicylique. Ce dernier est ensuite oxydé en catéchol ou en acide gentisique selon le type de bactéries. C'est là que commence la «voie centrale» qui convertit en plusieurs étapes l'intermédiaire catécholique en acétyl co-enzyme A et en acide pyruvique ou fumarique, des intermédiaires du cycle de Krebs.

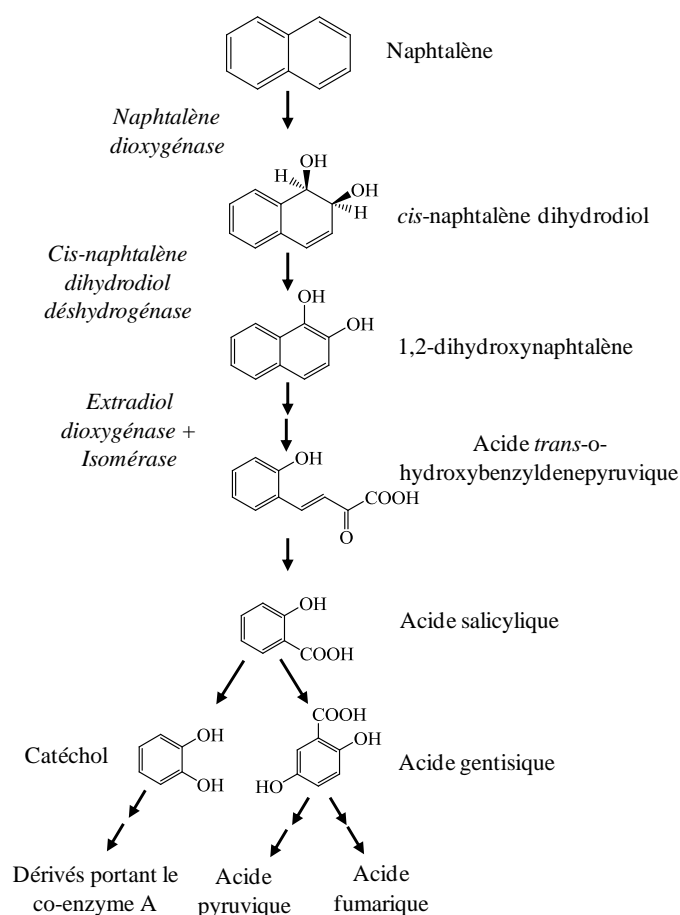


Figure I.12 : Etapes principales de la voie de dégradation du naphthalène par les bactéries aérobies. Inspiré de (Peng *et al.*, 2008).

2.5.1.2 Dégradation du phénanthrène par les bactéries

Les voies de dégradations aérobies du phénanthrène sont décrites Figure I.13 et Tableau I.3.

L'attaque initiale de la molécule tricyclique est réalisée par une dioxygénase qui, selon le type d'enzyme, incorpore les deux atomes de l'oxygène en positions 1,2-, 3,4- ou 9,10-. L'isomère 9,10, parfois produit par certaines bactéries du genre *Mycobacterium* par exemple (Moody *et*

al., 2001) est soit excrété soit converti en acide diphénique non métabolisable. La voie de dégradation la plus fréquente est celle qui commence par une dioxygénation en position 3,4. Les isomères hydroxylés en positions 1,2 et 3,4 sont oxydés par une déshydrogénase, ce qui permet le clivage du noyau aromatique par une extradiol-dioxygénase (coupure meta). L'intermédiaire formé est ensuite converti en 2 étapes en 2-hydroxy-1-naphtoate qui, chez les bactéries Gram négatif, donne le 1,2-dihydroxynaphtalène. Celui-ci est alors dégradé selon les mêmes étapes que celles de la voie de dégradation du naphtalène. Chez les bactéries Gram positif, le 2-hydroxy-1-naphtoate est le substrat d'une dioxygénase spécifique qui clive le noyau entre les carbones substitués par les groupements carboxyle et hydroxyle (Iwabuchi and Harayama, 1998).

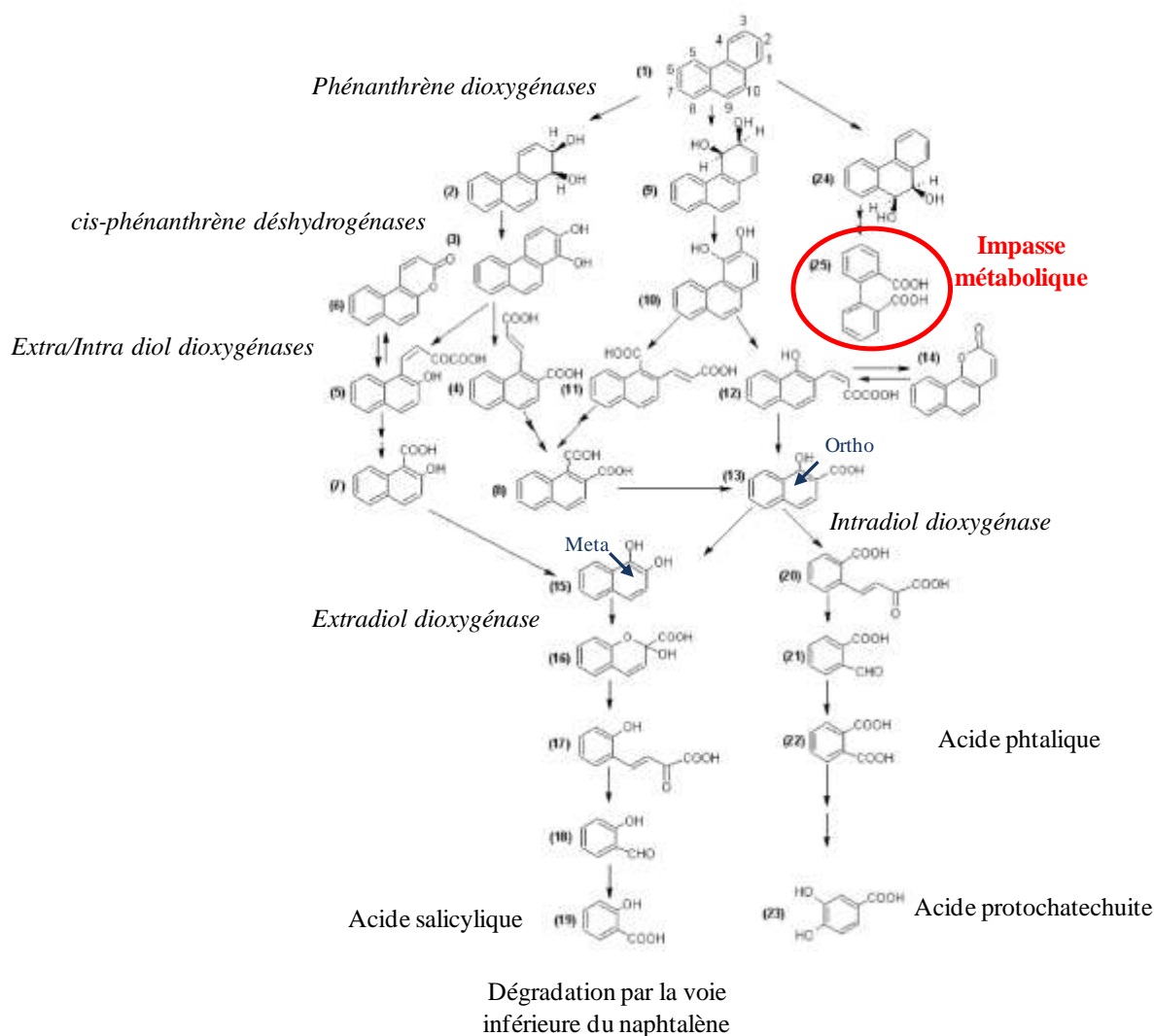


Figure I.13 : Voies de dégradation du phénanthrène par les bactéries aérobies (Peng *et al.*, 2008).

Les étapes suivantes conduisent à l'acide phtalique, puis à l'acide protocatéchuïque, par une voie semblable à celle qui métabolise le phénanthrène 1,2-dihydrodiol par coupure ortho (Figure I.13).

Tableau I.3 : Intermédiaires de dégradation du phénanthrène selon le mode d'attaque initiale Figure I.13

Voie métabolique supérieure de dégradation du phénanthrène (1)		
Voie 1,2	Voie 3,4	Voie 9,10
(2) <i>cis</i> -1,2-dihydroxy-1,2-dihydrophénanthrène	(9) <i>cis</i> -3,4-dihydroxy-3,4-dihydrophénanthrène	(24) <i>cis</i> -9,10-dihydroxy-1,2-dihydrophénanthrène
(3) 1,2-dihydroxyphénanthrène	(10) 3,4-dihydroxyphénanthrène	(25) acide 2,2'-diphenique
(4) acide 2-[(E)-2-carboxyvinyl]-1-naphtoïque	(11) acide 1-[(E)-2-carboxyvinyl]-2-naphtoïque	
(8) acide naphthalène-1,2-carboxylique		
(5) acide <i>trans</i> -4-(2-hydroxynaph-1-yl)-2-oxobut-3-énoïque	(12) acide <i>trans</i> -4-(1-hydroxynaph-2-yl)-2-oxobut-3-énoïque	
(6) 5,6-benzocoumarine	(13) acide 1-hydroxy-2-naphtoïque	
(7) acide 2-hydroxy-1-naphtoïque	(14) 7,8-benzocoumarine	
Voie métabolique inférieure de dégradation du phénanthrène (1)		
Voie Salicylate	Voie Protocatechuate	
(15) 1,2-dihydroxynaphtalène	(20) acide <i>trans</i> -2-carboxybenzalpyruvique	
(16) acide 2-hydroxy-2H-chromène-2-carboxylique	(21) carboxybenzaldéhyde	
(17) acide <i>trans</i> -o-hydroxybenzalpyruvique	(22) acide o-phtalique	
(18) salicylaldéhyde	(23) acide protocatechuïque	
(19) acide salicylique		

Lorsque les bactéries aérobies sont capables de les métaboliser, les HAP possédant plus de deux cycles aromatiques sont dégradés selon un schéma similaire : les noyaux aromatiques sont attaqués par une dioxygénase puis clivés l'un après l'autre. Cependant certaines étapes restent encore mal connues tant au plan du mécanisme que des enzymes mises en jeu (Kim *et al.*, 2007; Krivobok *et al.*, 2003; Kweon *et al.*, 2007; Peng *et al.*, 2008).

Certaines voies métaboliques des HAP produisent des intermédiaires non métabolisables (Ex : (25) de la voie du phénanthrène). Ils sont souvent plus hydrophiles que le produit de départ, plus réactifs et parfois plus toxiques. Certains intermédiaires de type catéchol peuvent inactiver les enzymes extradiol dioxygénases, par oxydation de leur site actif (Jouanneau, 2010). L'accumulation de composés catécholiques peut conduire à la formation d'espèces réactives de l'oxygène et provoquer des dommages aux protéines ou à l'ADN.

2.5.1.3 Organisation des opérons cataboliques chez les bactéries

Les gènes cataboliques de dégradation des HAP sont souvent regroupés en opérons, sur le chromosome ou plus fréquemment sur des éléments mobiles tels que des plasmides ou des transposons. Chez les bactéries du genre *Pseudomonas*, les gènes cataboliques du naphthalène (gènes *nah* ou *nag*) sont organisés en deux opérons. Un premier opéron code les enzymes de la voie supérieure (*nah*) qui convertissent le naphthalène en salicylate. Un second opéron code les enzymes de la voie inférieure (*sal*) qui convertissent le salicylate en pyruvate et en acetyl Co-A. Ces deux opérons sont régulés par une protéine activatrice de la transcription, codée par *nahR* (Peng *et al.*, 2008). Les opérons les plus étudiés sont ceux portés par les plasmides NAH7 de *P. putida* G7 (gènes *nah*) (Kurkela *et al.*, 1988), DTG1 de *P. putida* NCIB9816-4 (gènes *ndo*) (Simon *et al.*, 1993), C18 de *Pseudomonas* sp C18 (gènes *dox*) (Denome *et al.*, 1993). En plus des gènes cataboliques, l'opéron de la voie supérieure porte parfois un gène impliqué dans le processus de chimiotactisme (gène *nahY*), comme cela a été mis en évidence chez *P. putida* G7 (Grimm and Harwood, 1999).

Cependant, les opérons du catabolisme du naphthalène et du phénanthrène ne suivent pas tous ce schéma. Ainsi, l'opéron portant la phénanthrène-dioxygénase de *Burkholderia* sp RP007 est situé sur un plasmide mais les gènes des transporteurs d'électrons nécessaires au fonctionnement de la dioxygénase ne sont pas présents. La fonction serait assurée soit par des gènes de transporteurs d'électrons situés à un autre endroit du génome de RP007, soit par des gènes non spécifiques (Laurie and Lloyd-Jones, 1999). De même, l'opéron portant la dioxygénase NagAcAd, de *Ralstonia* sp. U2 et situé sur un plasmide, porte deux gènes supplémentaires *nagGH* qui codent une salicylate-5-hydroxylase catalysant la conversion du salicylate en gentisate (Fuenmayor *et al.*, 1998).

A l'inverse, l'opéron catabolique de *Acidovorax* sp. NA3, portant la dioxygénase PhnAcAd (Singleton *et al.*, 2009), et l'opéron catabolique de *Polaromonas* sp. CJ2, portant la dioxygénase NagAcAd (Jeon *et al.*, 2006), se situent sur l'ADN chromosomique.

Chez les bactéries Gram⁺, comme *Nocardioides* sp. KP7 qui dégrade le phénanthrène, l'opéron catabolique portant la dioxygénase PhdAB est aussi situé sur le chromosome (Saito *et al.*, 2000).

Les gènes des Sphingomonadaceae et des Mycobactéries sont, quant à eux, beaucoup plus dispersés dans l'ADN génomique (Pinyakong *et al.*, 2003a). Ainsi *Mycobacterium vanbaalenii* PYR-1 et *Mycobacterium* 6PY1, possèdent plusieurs gènes de dioxygénases, *nidBA*, *nidA3B3* et *pdoA2B2* qui sont éloignés les uns des autres et des gènes codant les transporteurs d'électrons associés (Khan *et al.*, 2001; Kim *et al.*, 2006; Krivobok *et al.*, 2003). Chez *Sphingomonas aromaticivorans* F199, les gènes *bph* de dégradation des HAP sont intercalés avec des gènes *xyl* de dégradation des monoaromatiques sur un megaplasmide de 189 kbp (pNL1) (Romine *et al.*, 1999).

2.5.1.4 Dégradation des HAP par les champignons

Les champignons du sol de type micromycètes sont capables de dégrader les composés aromatiques suivant deux types de voies métaboliques (Cerniglia, 1992; Peng *et al.*, 2008).

Les champignons lignolytiques excrètent trois types d'enzymes solubles, les lignine-peroxydases, les manganèse-peroxydases, ou les phénol-oxydases, ainsi que des enzymes produisant du peroxyde d'hydrogène. Ces enzymes sont impliquées dans la dégradation de la lignine, un polymère constitué de structures phénoliques (Haritash and Kaushik, 2009). Ayant une faible spécificité du substrat, elles attaquent aussi les HAP, y compris quand ils sont piégés dans la matrice du sol. L'attaque de l'un des noyaux aromatiques libère un composé hydroxylé, puis conduit à la production de quinones par oxydation radicalaire (Figure I.14).

Les champignons non-lignolytiques utilisent des mono-oxygénases de type cytochrome P450 pour insérer un atome d'oxygène dans un cycle aromatique et former un époxyde très réactif. Celui-ci réagit avec l'eau pour former un aromatique hydroxylé. Les arènes hydroxylés sont souvent oxydés une seconde fois pour former des *trans*-dihydrodiols. Lorsqu'elle est partielle, la dégradation des HAP par les champignons génère des composés phénoliques, des tétralones, des quinones, ou des diol-époxydes eux-mêmes très réactifs et qui peuvent être toxiques (Cerniglia, 1997; Meulenberg *et al.*, 1997; Sutherland, 1992).

La majorité des champignons est capable d'attaquer les HAP de haut poids moléculaire produisant des intermédiaires plus solubles et plus réactifs, qui sont potentiellement métabolisables par certaines bactéries aérobies (Boonchan *et al.*, 2000; Cerniglia, 1997).

2.5.1.5 Dégradation des HAP par les microalgues

Plusieurs études ont rapporté l'implication des microalgues procaryotes et eucaryotes dans la dégradation des HAP (Cerniglia *et al.*, 1980; Narro *et al.*, 1992; Tam *et al.*, 2007). Ces microorganismes présentent plusieurs voies d'attaque des cycles aromatiques (Figure I.14), notamment *via* des mono-oxygénases, suivant une voie similaire à celle des champignons, ou *via* des dioxygénases comme le rapportent les travaux de Warshawsky (Warshawsky *et al.*, 1995). Ces microorganismes, principalement photoautotrophes, dégradent les composés aromatiques pour produire des intermédiaires utilisables pour la biosynthèse de macromolécules (Gibson and Harwood, 2002). De même, les microalgues et les bactéries s'associent au sein de consortiums dans lesquels, les uns produisent de l'oxygène par la photosynthèse et les autres dégradent les HAP pour produire du CO₂. Cette synergie renforce le potentiel de dégradation des HAP (Borde *et al.*, 2003).

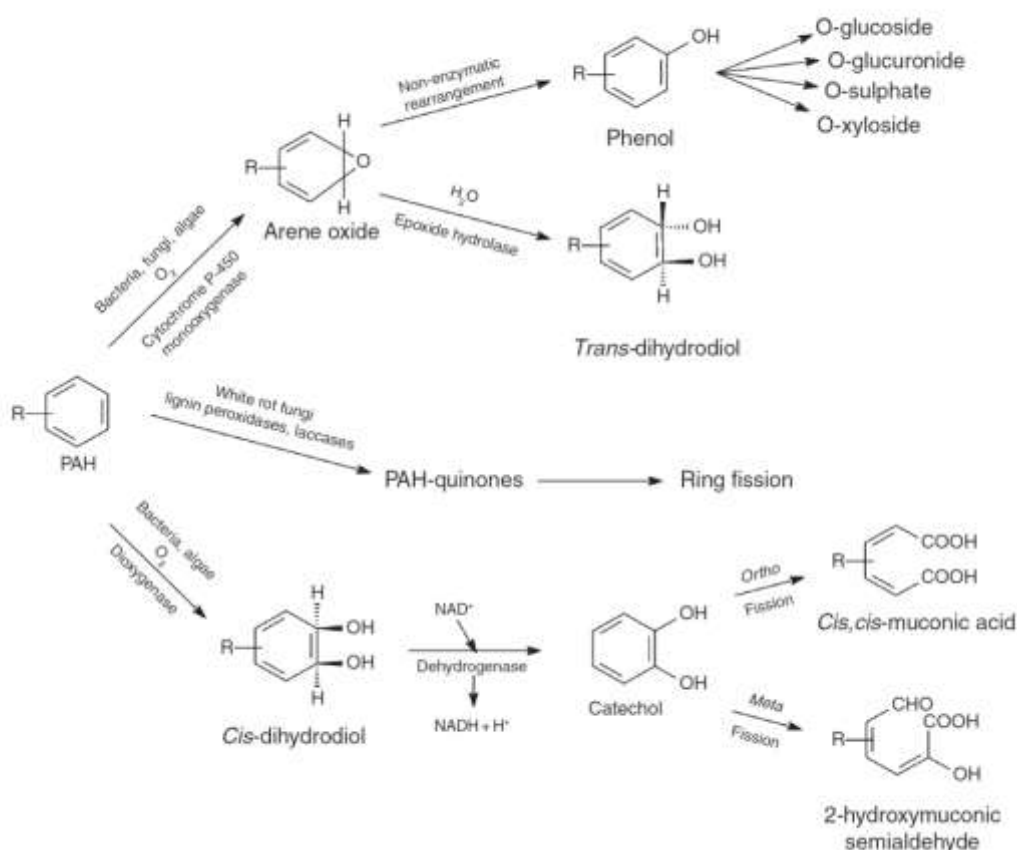


Figure I.14 : Différentes voies de dégradation aérobie des HAP selon le type de microorganismes (Cerniglia, 1992).

2.5.2 Voies de dégradation anaérobies chez les bactéries

En absence d'oxygène, les bactéries métabolisent les HAP par voie anaérobie, mais beaucoup plus lentement qu'en aérobiose. Peu d'études décrivent les voies métaboliques impliquées dans cette dégradation. Les microorganismes utilisent majoritairement le nitrate et le sulfate comme accepteurs d'électrons et, dans une moindre mesure, le fer ferrique, le chlorate, le manganèse ou le CO₂ (Gibson and Harwood, 2002; Meckenstock *et al.*, 2004). La dégradation du méthylnaphtalène, qui est la plus documentée, est associée à la réduction du sulfate en sulfure ou du nitrate en azote gazeux. L'attaque initiale est une addition de la molécule de fumarate sur le méthyl pour donner le naphthyl-2-méthyl-succinate. Une molécule de coenzyme A (CoA) est ensuite transférée du succinyl-CoA au naphthyl-2-méthyl-succinate. Le carbone méthylé du composé obtenu est ensuite beta-oxydé. Puis une molécule d'H₂O serait ajoutée, suivi d'une étape d'oxydation et d'une étape de clivage de la chaîne latérale, conduisant à deux intermédiaires : l'acide 2-naphtoïque et le succinyl-CoA (Annweiler *et al.*, 2000; Meckenstock *et al.*, 2004).

Le naphthalène, quant à lui, serait soit activé par carboxylation (Meckenstock *et al.*, 2000), soit méthylé pour produire le même intermédiaire, l'acide 2-naphtoïque (Safinowski and Meckenstock, 2006).

L'acide 2-naphtoïque est réduit par étapes successives pour donner le decalin-2-carboxylate (Widdel and Rabus, 2001). Par la suite, l'ouverture du noyau conduirait à la production d'un intermédiaire de la voie centrale de dégradation dont un composé cyclohexanoïque serait un intermédiaire (Meckenstock *et al.*, 2004).

Enfin, l'étape initiale de la dégradation anaérobie du phénanthrène associée à la réduction du sulfate serait une activation par carboxylation directe (Suflita *et al.*, 2007).

De manière générale, l'utilisation de substrats radio-marqués ou marqués au ¹³C montre la production de ¹⁴CO₂ ainsi que l'incorporation de carbone marqué dans la biomasse, suggérant que la minéralisation complète du naphthalène, du phénanthrène et du fluoranthène est possible en conditions anoxiques (Meckenstock *et al.*, 2000; Rockne *et al.*, 2000). La dégradation des

HAP en culture pure ou dans des sédiments est un processus lent qui peut durer de quelques semaines à plusieurs mois (Coates *et al.*, 1997; Meckenstock *et al.*, 2000; Suflita *et al.*, 2007).

Bien que la dégradation anaérobie des HAP soit plus lente qu'en présence d'oxygène, l'énergie produite par leur oxydation avec des nitrates, ou du fer ferrique comme accepteurs est proche de celle produite par oxydation en présence d'O₂. En revanche, l'oxydation couplée à la réduction des sulfates est nettement moins favorable en termes de production d'énergie. (Diaz, 2004)

2.6 La diversité microbienne dégradant les HAP dans les sols

Les voies de dégradation décrites précédemment ont été caractérisées grâce à l'isolement de microorganismes sur des milieux sélectifs contenant un HAP particulier comme unique source de carbone. Les premières bactéries étudiées, isolées sur naphtalène, appartenaient au genre *Pseudomonas* (Grund and Gunsalus, 1983; Kurkela *et al.*, 1988; Simon *et al.*, 1993; Yang *et al.*, 1994). Par la suite, d'autres bactéries appartenant aux genres *Sphingomonas* et *Mycobacterium* ont largement été étudiées après isolement sur différents HAP (Kim *et al.*, 2007; Krivobok *et al.*, 2003; Shi *et al.*, 2001; Willison, 2004). Plus récemment, des méthodes indépendantes de la culture ont permis d'enrichir nos connaissances sur des microorganismes actifs dans les sols contaminés aux HAP (Jeon *et al.*, 2006; Ono *et al.*, 2007; Singleton *et al.*, 2009).

2.6.1 Les Gammaprotéobactéries

Des bactéries du genre *Pseudomonas*, appartenant à la classe des Gammaprotéobactéries, ont servi de modèles pour étudier le métabolisme du naphtalène et du phénanthrène. Les opérons cataboliques de *P. putida* G7 et *P. stutzeri* AN10 ont été caractérisés après isolement sur naphtalène (Bosch *et al.*, 1999; Menn *et al.*, 1993). De même, *P. aeruginosa* a été isolée pour ses capacités à dégrader le phénanthrène (Romero *et al.*, 1998). *P. putida*, et *P. aeruginosa* sont aussi connues pour leur faculté de chimiotactisme vis à vis des HAP (Grimm and Harwood, 1999). D'autres *Pseudomonas* ont été détectées dans des sols contaminés par une pollution diffuse (Johnsen *et al.*, 2002). Une autre Gammaprotéobactérie, *Stenotrophomonas*,

est capable de dégrader le phénanthrène ainsi que des HAP de haut poids moléculaire comme le benz[a]anthracène, le benzo[a]pyrène, le fluoranthène et le pyrène (Juhasz *et al.*, 2000).

En milieu marin, des bactéries du genre *Cycloclasticus*, d'autres Gammaprotéobactéries, ont été identifiées comme acteurs de la dégradation du naphthalène et du phénanthrène notamment (Geiselbrecht *et al.*, 1996).

2.6.2 Les Alphaprotéobactéries

Les Sphingomonades sont des microorganismes capables de dégrader un large spectre de composés aromatiques (Leys *et al.*, 2004). Ils ont été isolés ou détectés dans des sols fortement contaminés aux hydrocarbures. Les Sphingomonadaceae comprennent plusieurs genres appelés *Sphingomonas*, *Sphingobium*, *Novosphingobium*, et *Sphingopyxis* (Bastiaens *et al.*, 2000). *Sphingobium yanoikuyae* sp. B1 (anciennement *Beijerinckia* sp. B1) a été isolé pour ses capacités à utiliser le naphthalène comme unique source de carbone et d'énergie (Zylstra *et al.*, 1997). *Sphingomonas* sp. LH-128 a été isolé sur le phénanthrène (Schuler *et al.*, 2009), tout comme *Sphingomonas paucimobilis* TNE12 (Shuttleworth and Cerniglia, 1996). *Sphingomonas* sp. CHY-1, bien que préférant le naphthalène, a été isolé sur chrysène (Willison, 2004). Certains Sphingomonades, tels que *Sphingomonas paucimobilis* EPA 505, dégradent le fluoranthène et co-métabolisent le benzo[a]pyrène (Ye *et al.*, 1996).

2.6.3 Les Betaprotéobactéries

Plusieurs espèces de Betaprotéobactéries sont aussi capables de se développer sur naphthalène ou phénanthrène (Johnsen *et al.*, 2002). La famille des Burkholderiales est la plus représentée à ce jour.

Burkholderia sp. C3 (Seo *et al.*, 2007) et *Burkholderia* sp. RP007 (Laurie and Lloyd-Jones, 1999) ont été isolées sur phénanthrène. La souche RP007 est aussi capable de minéraliser le naphthalène et l'anthracène. *Ralstonia* sp. U2, isolée par enrichissement sur naphthalène, (Fuenmayor *et al.*, 1998) métabolise ce HAP par une voie particulière impliquant le gentisate comme intermédiaire. *Polaromonas naphthalenivorans* CJ2, possédant des gènes cataboliques similaires à ceux de *Ralstonia* sp. U2, minéralise le naphthalène (Jeon *et al.*, 2006).

Alcaligenes sp. AKF2 dégrade le phénanthrène par la voie du phtalate mais ne minéralise pas le naphthalène (Kiyohara *et al.*, 1982). Bien que les gènes cataboliques de *Acidovorax* sp. NA3 soient très proches de ceux de la souche AKF2, cette bactérie minéralise les HAP de 2-3 cycles tels que le naphthalène ou le phénanthrène, ainsi que certains HAP de haut poids moléculaire tels que le chrysène, le benz[a]anthracène ou le benzo[a]pyrène (Singleton *et al.*, 2009).

2.6.4 Les Actinobactéries

Les bactéries Gram⁺, et plus particulièrement celles appartenant au phylum des Actinobactéries, sont aussi impliquées dans la dégradation des HAP. Les mycobactéries, genre le plus étudié, sont capables de se développer en utilisant le phénanthrène, l'anthracène ou le pyrène comme uniques sources de carbone et sont aussi impliquées dans la dégradation d'HAP de haut poids moléculaire, comme le fluoranthène et le benzo[a]pyrène. Ces microorganismes ont été détectés dans des environnements à pollution diffuse, avec une faible concentration et une faible biodisponibilité en HAP. Ils produisent dans leur paroi externe des acides mycoliques, acides gras alkylés et hydroxylés, qui forment une couche quasi cristalline protégeant la cellule de la toxicité des polluants. La présence de cette barrière de protection associée à la fluidité de leur bicouche lipidique sert aussi à l'internalisation des HAP (Jouanneau *et al.*, 2011; Leys *et al.*, 2005).

Mycobacterium sp. 6PY1, isolée sur pyrène, est aussi capable de dégrader le phénanthrène et le fluoranthène (Krivobok *et al.*, 2003). *Mycobacterium* sp. RJGII-135 est capable de dégrader le pyrène, le benzo[a]pyrène et le benz[a]anthracène (Grosser *et al.*, 1991; Schneider *et al.*, 1996). La voie de dégradation du pyrène de *Mycobacterium* sp. PYR-1 a été particulièrement documentée (Heitkamp *et al.*, 1988; Kim *et al.*, 2008; Kim *et al.*, 2007; Rafii *et al.*, 1992; Stingley *et al.*, 2004).

D'autres actinobactéries, du genre *Arthrobacter* ou *Nocardioides*, sont capables de dégrader les HAP de faible poids moléculaire tels que le naphthalène, le phénanthrène, l'anthracène, le fluorène. *Nocardioides* sp. KP7 dégrade le phénanthrène par la voie phtalate, mais pas le naphthalène (Saito *et al.*, 2000). *Janibacter* sp. YY1, isolé sur dibenzofurane peut utiliser le fluorène et le dibenzothiophène comme sources uniques de carbone et d'énergie, et peut co-

métaboliser le naphthalène et le phénanthrène (Yamazoe *et al.*, 2004). La souche *Rhodococcus* sp. UW1, isolée sur pyrène, utilise, quant à elle, le phénanthrène, l'anthracène, le fluoranthène et le chrysène comme sources uniques de carbone et d'énergie. Cette souche est aussi capable de co-métaboliser le naphthalène et le fluorène notamment (Walter *et al.*, 1991).

2.6.5 Identification d'acteurs de la dégradation des HAP par une approche indépendante de la culture

Singleton et collaborateurs ont appliqué une approche culture-indépendante (SIP) pour identifier les espèces bactériennes impliquées dans la dégradation du naphthalène, du phénanthrène et du pyrène dans un bioréacteur contenant des boues contaminées aux HAP. Ainsi, des espèces apparentées aux genres *Pseudomonas* et *Ralstonia* seraient impliquées dans la dégradation du naphthalène, alors que des espèces du genre *Acidovorax* et des Betaprotéobactéries ont été identifiées sur phénanthrène (Singleton *et al.*, 2005). Des bactéries inconnues jusque là et appartenant aux Alpha, aux Beta et aux Gamma protéobactéries sont impliquées dans le métabolisme du pyrène (Singleton *et al.*, 2006). Enfin, quand le pyrène et le phénanthrène sont simultanément présents, *Acidovorax* métabolise le phénanthrène, mais pas le pyrène. A l'inverse, l'OTU de Gammaprotéobactéries détectées sur pyrène dégrade dans une moindre mesure le phénanthrène. La présence de ce dernier diminue, en outre, leur capacité à dégrader le pyrène. (Singleton *et al.*, 2007).

2.6.6 Les acteurs de la dégradation anaérobie des HAP

Des travaux sur la dégradation anaérobie des HAP ont mis en évidence des bactéries membres de la classe des Gammaprotéobactéries, proches parents de *Pseudomonas stutzeria* ou *Vibrio pelagus*, capables de dégrader le naphthalène en conditions dénitrifiantes (Rockne *et al.*, 2000) ou des Deltaprotéobactéries sulfato-réductrices de la famille *Desulfobacteraceae*, dégradant le phénanthrène (Suflita *et al.*, 2007).

3. Les arène-dioxygénases

3.1 Organisation structurale

Les arène-dioxygénases (ring-hydroxylating dioxygenases ou RHDs) sont des métallo-enzymes très répandues chez les bactéries. Elles catalysent l'oxydation initiale de composés aromatiques variés tels que les HAP. Comme cela a été présenté précédemment, les RHD catalysent l'hydroxylation simultanée de deux carbones adjacents de l'un des noyaux aromatiques, générant des *cis*-dihydrodiols. Ces enzymes multicomposantes associent une oxydoréductase de type flavoprotéine, une ferrédoxine et une oxygénase dont le site actif interagit avec le composé aromatique. Certaines dioxygénases ne sont formées que de deux composantes, une réductase et une oxygénase.

3.1.1 Les transporteurs d'électrons

Les réductases sont des protéines de ~ 45 kDa classées en deux familles. Une première famille contient deux groupements prosthétiques, un groupement de type flavine (FAD ou FMN), et un centre [2Fe-2S] de type plante, dans lequel les atomes de fer sont coordonnés par quatre résidus cystéine (Ferraro *et al.*, 2005). La réductase de la phtalate dioxygénase de *Pseudomonas cepacia* sp. PHK en est un exemple (Correll *et al.*, 1992). La deuxième famille ne possède qu'un groupement de type FAD, comme par exemple la réductase de la biphényle-dioxygénase de *Pseudomonas* sp. KKS102 (Senda *et al.*, 2000).

Les ferrédoxines sont des petites protéines (~ 12 kDa) comportant un centre [2Fe-2S] de type plante ou bien de type Rieske [2Fe-2S]. Dans le cluster Rieske, les atomes de fer sont coordonnés par deux résidus cystéine et deux résidus histidine (Figure I.15). Dans certains cas, la ferrédoxine possède un centre [3Fe-4S] ou [4Fe-4S] comme cela a été trouvé dans certaines dioxygénases oxydant les HAP, notamment chez les Actinobactéries telles que *Nocardioïdes* sp. KP7 (Saito *et al.*, 2000).

3.1.2 L'oxygénase terminale et son mécanisme catalytique

La composante oxygénase est une protéine multimérique associant des sous-unités alpha (50 kDa) et beta (20 kDa) le plus souvent sous la forme d'un complexe $\alpha_3\beta_3$. La structure

cristallographique de la composante catalytique de la naphthalène-dioxygénase de *Pseudomonas* sp. NCIB 9816-4 (NDO) a une forme générale qui rappelle celle d'un champignon, avec les trois sous-unités α reposant sur une base formée des trois sous-unités β (Kauppi *et al.*, 1998). La carbazole dioxygénase de *Pseudomonas resinovorans* CA10 fait un peu exception en ce sens que sa composante catalytique n'est composée que de sous-unités alpha (Nam *et al.*, 2002). Pour cette raison, la plupart des spécialistes pensent que la sous-unité β a un rôle essentiellement structural (Ferraro *et al.*, 2005).

Chaque sous-unité α est constituée de deux domaines, un domaine N-terminal contenant un centre [2Fe-2S] de type Rieske et un domaine catalytique contenant un centre Fe(II) mononucléaire dans la partie C-terminale.

La sphère de coordination du centre [2Fe-2S] de NDO est déterminée par une séquence comprenant des résidus conservés et formant le motif suivant : **C**X**H**RGX₈GNX₅**C**X**Y**H**G**. Le Fe_1 est coordonné par deux histidines (His82 et His104) et passe de l'état ferrique (FeIII) à l'état ferreux (FeII) lors de la réduction du cluster. Le Fe_2 est quant à lui lié à deux cystéines (Cys80 et Cys100) et reste sous forme ferrique (Figure I.15).

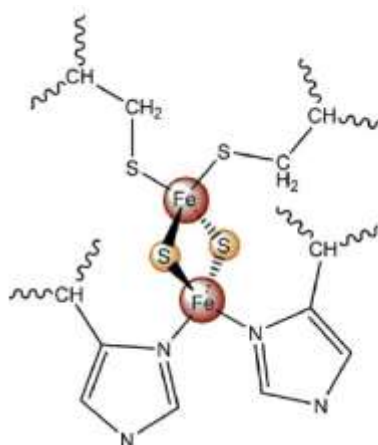


Figure I.15 : Schéma de la sphère de coordination du centre [2Fe-2S] des RHD (Ferraro *et al.*, 2005).

La poche de fixation du substrat offre un environnement hydrophobe, ouvert d'un côté sur le solvant et se terminant par l'atome de fer catalytique, de l'autre côté. Elle contient trois résidus conservés qui sont His208, His213 et Asp205 chez NDO. Les deux histidines sont avec Asp362 les trois ligands du fer catalytique. D'autres résidus semi-conservés bordent la poche catalytique de la naphthalène-dioxygénase NDO de *Pseudomonas* sp. NCIB 9816-4 :

(Asn201-Phe202-**His202**-His213-Phe358) (Carredano *et al.*, 2000). Les résidus His208-Val209, Ala206-Trp358 ainsi que Phe352 seraient impliqués dans l'interaction avec le substrat et contrôlèrent le positionnement de celui-ci par rapport à l'atome de fer du site actif. Ce contrôle de l'orientation du substrat dans le site actif détermine probablement la régio- et l'énantio-sélectivité de la réaction catalytique (Parales *et al.*, 2000; Parales and Resnick, 2006).

Pendant la catalyse, un transfert d'électrons a lieu entre le centre Rieske d'une sous-unité α et le fer catalytique de la sous-unité α adjacente qui est plus proche (12 Å) que celui de la même sous-unité (44 Å). Les deux centres métalliques sont reliés par liaisons hydrogène impliquant un acide aspartique (Asp205) et deux ligands histidine (His208 et His104) (Ferraro *et al.*, 2005). Cet acide aspartique situé à l'interface entre les deux sous-unités α constitue un pont à électrons plausible entre le centre Rieske et le fer catalytique. La réduction du centre [2Fe-2S] s'accompagne d'une protonation du résidu His104, créant un pont hydrogène. Alors qu'à l'état oxydé, le fer catalytique est lié à deux histidines, un aspartate et une molécule d'eau, la réduction induit un changement de conformation entraînant la liaison à une seconde molécule d'eau ; le fer est alors dans une conformation octaédrique. Cette géométrie impliquant cinq ligands rend le fer catalytique plus apte à réagir avec l'oxygène moléculaire.

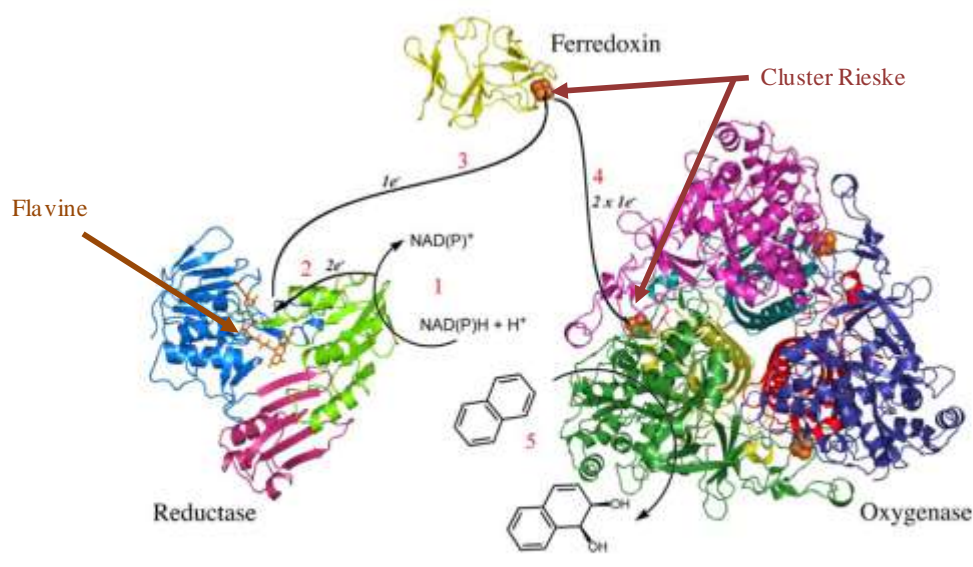


Figure I.16 : La réductase oxyde (1) NAD(P)H en NAD+ au niveau de son site de fixation NAD(P)H et obtient deux électrons (2). Les électrons sont stockés dans la flavine jusqu'au transfert vers la ferrédoxine (3). La ferrédoxine transfère un par un les électrons reçus vers le centre Rieske de l'oxygénase (4). Oxydation du substrat au niveau du Fe(II) catalytique (5). (Ferraro *et al.*, 2005).

Lors du cycle de catalyse, la réductase oxyde le NADH qui libère alors deux électrons. Ils sont transférés un par un *via* le groupement FAD au centre fer/soufre de la ferrédoxine, et de là au centre [2Fe-2S] de la sous unité α de l'oxygénase. Le centre Rieske réduit, transfère à son tour un électron à l'atome de fer du site catalytique qui fixe le substrat (Figure I.16). La protéine changerait alors de conformation pour laisser entrer l'oxygène moléculaire afin de l'activer *via* la formation d'un intermédiaire Fe(III) (hydro) peroxo (Kovaleva and Lipscomb, 2008; Wolfe and Lipscomb, 2003). Contrairement au mode d'action classique des autres oxygénases connues selon lequel l'oxygène se fixe au fer en position distale, ici, la molécule d'oxygène se fixe en position latérale (Karlsson *et al.*, 2003). Ce positionnement fait que les deux atomes de l'oxygène sont proches des 2 atomes de carbone du substrat aromatique, ce qui facilite leur hydroxylation simultanée (Figure I.17). Il a été montré que l'oxygénase à elle seule pouvait catalyser l'oxydation du substrat sans transporteurs d'électrons à condition d'avoir été réduite au préalable (Wolfe *et al.*, 2001). Cependant, la libération du dihydrodiol formé nécessite la réduction du fer catalytique, et donc l'apport d'un électron supplémentaire.

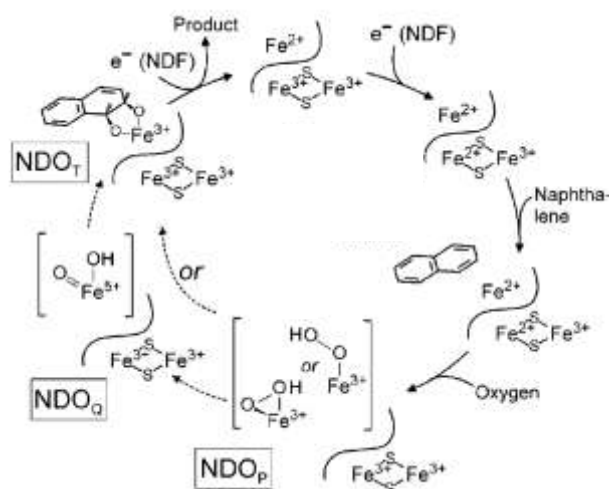


Figure I.17 : Mécanisme d'action proposé pour l'oxygénation du naphtalène par NDO de *Pseudomonas putida* NCIB 9816-4 (Wolfe and Lipscomb, 2003).

3.2 La dioxygénase de *Sphingomonas* sp. CHY1

Sphingomonas sp. CHY1 possède une dioxygénase à trois composantes capable d'initier l'oxydation de huit des seize HAP majoritaires, y compris les composés de quatre ou cinq cycles (Demaneche *et al.*, 2004; Jouanneau *et al.*, 2006). La structure tridimensionnelle de la composante catalytique, appelée PhnI, a été résolue en 2007 (Jakoncic *et al.*, 2007a; Jakoncic *et al.*, 2007b). Celle-ci est globalement comparable à celle de NDO, et diffère notamment par la forme de la poche catalytique (Figure I.18).

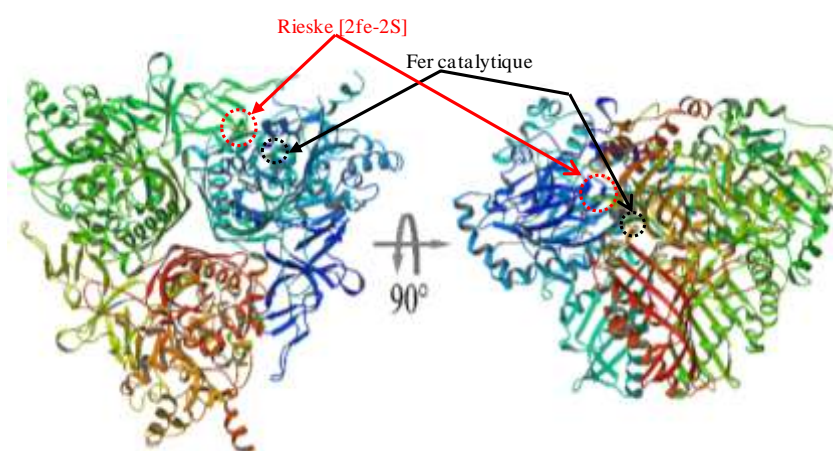


Figure I.18 : Structure cristallographique de PhnI représentant le complexe hexamérique $\alpha_3\beta_3$. Les trois hétérodimères $\alpha\beta$ sont représentés en rouge, vert et bleu, avec les sous-unités β en plus clair. La position relative des centres Rieske et des centres à fer de deux sous-unités adjacentes est indiquée par les flèches. (Jakoncic *et al.*, 2007b)

3.2.2 Eléments structuraux de la sous-unité α

Le centre Rieske de la sous-unité α présente une structure similaire à celle de NDO : le Fe_1 est coordonné par les résidus His82 et His103 et le Fe_2 est coordonné par les résidus Cys80 et Cys100 (Figure I.19). En revanche, la poche hydrophobe de fixation du substrat, est beaucoup plus volumineuse (91 \AA^3) que celle des dioxygénases caractérisées jusqu'alors : 45 \AA^3 pour la biphényl dioxygénases de *Rhodococcus* sp. RHA1, et 65 \AA^3 pour la naphthalène-dioxygénase de *Pseudomonas putida* NCIB9816-4 (NDO). L'accès à cette poche est délimité par deux boucles flexibles LI et LII, exposées au solvant, dont l'une (LI) contribue à l'allongement d'environ 2 \AA de la poche.

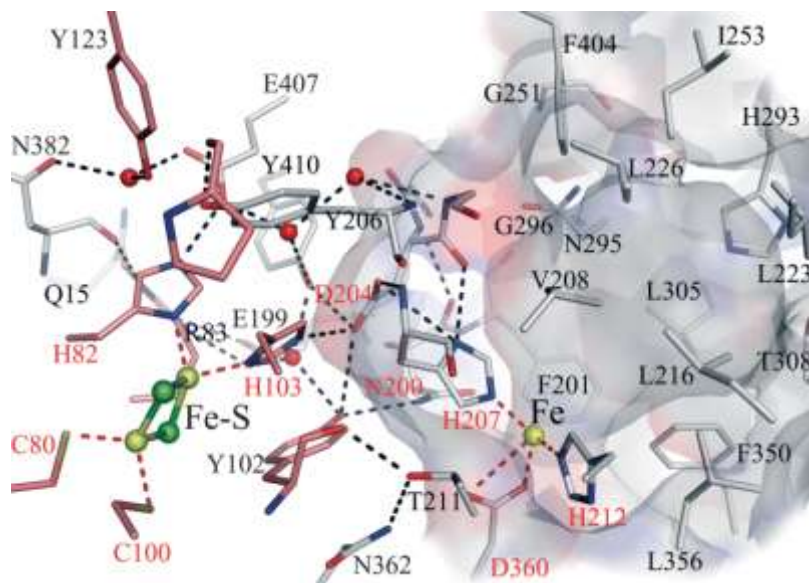


Figure I.19 : Vue partielle de la structure de PhnI montrant le domaine Rieske et le domaine catalytique de deux sous-unités α adjacentes. Les résidus impliqués dans les sphères de coordination des atomes de fer ainsi que le résidu Asp204, qui serait impliqué dans le transfert d'électrons entre les 2 centres métalliques sont notés en rouge. Les molécules d'eau impliquées dans le réseau de liaisons hydrogène sont représentées par des sphères rouges. En fond, la surface de la poche catalytique représente l'espace interne disponible pour le substrat (Jakoncic *et al.*, 2007b).

On peut distinguer trois régions de cette poche selon leur distance du fer catalytique.

Les résidus Phe350, Phe404, Leu356 Gly205, Val208 et Thr308 de la région centrale contribuent à l'élargissement de la poche catalytique de PhnI par rapport à NDO. Plus particulièrement, le résidu Phe350 serait impliqué dans la sélection du substrat et la régio-spécificité de l'attaque des HAP (Jakoncic *et al.*, 2007b).

Au niveau de la région distale, les résidus Leu223 et Leu226 appartenant à la boucle LI et les résidus Ile253 et Ile260 de la boucle LII, contrôleraient l'accès à la poche catalytique et auraient une influence dans la sélection du substrat (Jakoncic *et al.*, 2007b).

Tout comme dans NDO, le résidu Asp204 assure la jonction entre le centre Rieske d'une sous-unité α et le fer catalytique de la sous-unité adjacente.

3.2.3 Spécificité du substrat

La poche de fixation du substrat se présente sous une forme trapézoïdale plus étroite dans la région proximale ce qui lui permet d'accueillir un grand nombre de substrats mais aussi de les orienter, contribuant sans doute à la régio- et stéréo-spécificité de la réaction. Ainsi, des modélisations indiquent que les HAP à 2 cycles seraient bloqués dans une position unique en

accord avec l'attaque du naphthalène en positions 1, 2 et celle du biphenyl en positions 2,3 (Jakoncic *et al.*, 2007a).

A cause de sa forme angulaire qui induit des contraintes stériques, la modélisation prédit que le phénanthrène adopte une seule position compatible avec une hydroxylation des carbones 3 et 4. Cette prédiction est conforme à la régiospécificité trouvée expérimentalement puisque le *cis*-3,4-dihydrodiol est le seul produit de la réaction (Jouanneau *et al.*, 2006).

L'enzyme génère plusieurs produits mono- ou dihydroxylés avec le fluorène et le fluoranthène, suggérant que ces substrats se fixent au site actif dans plusieurs positions possibles. L'oxydation du chrysène génère un produit majoritaire, le *cis*-3,4-dihydrodiol. Un second produit tetrahydroxylé, le bis-*cis*-3,4,9,10-dihydrodiol, a aussi été détecté. Il résulte d'une seconde dioxygénation du diol.

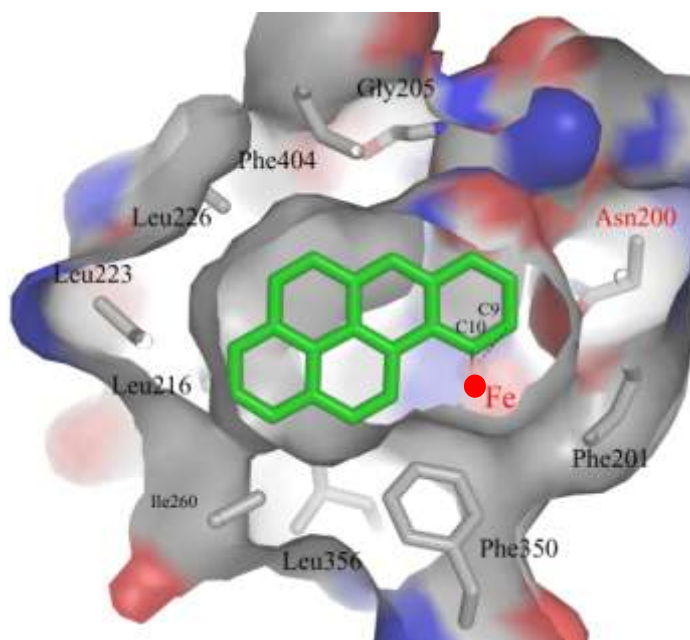


Figure I.20 : Modélisation du benzo[a]pyrène dans la poche catalytique de PhnI de *Sphingomonas* CHY-1 (Jakoncic *et al.*, 2007a)

Selon la modélisation structurale, le benzo[a]pyrène (BaP) ne peut se placer que dans une seule orientation offrant les positions 9,10 à l'hydroxylation (Figure I.20). Cette prédiction est confirmée par les dosages *in vitro* dans lesquels l'oxydation du BaP ne produit que le *cis*-9,10-dihydrodiol.

Enfin, le benz[a]anthracène (BaA), HAP à 4 cycles, peut être attaqué en position 1,2 sans grand changement conformationnel et en position 10,11, moyennant un réarrangement des

résidus Leu223 et Phe300. Expérimentalement, l'oxydation du BaA génère trois dérivés possibles : le *cis*-1,2-dihydrodiol est le plus abondant, le *cis*-8,9-dihydrodiol et le *cis*-10,11-dihydrodiol. Un quatrième sous-produit tetrahydroxylé, le bis-*cis*-1,2,10,11-dihydrodiol, a été identifié. Il se forme par une seconde hydroxylation des 1,2-diol et 10,11-diol.

PhnI de *Sphingomonas* CHY-1 est la première dioxygénase décrite qui présente une aussi large gamme de substrats polyaromatiques hydroxylables (Jouanneau *et al.*, 2006).

3.3 Diversité des dioxygénases

Les dioxygénases ont été classées en quatre familles selon les similitudes de séquence de leur sous-unités alpha (Nam *et al.*, 2001). Il se trouve que ce classement regroupe les enzymes selon leur spécificité du substrat. La famille I comporte les dioxygénases de type α_3 , les benzotate et toluate dioxygénases sont regroupées dans la famille II, les dioxygénases oxydant le naphthalène et les HAP forment la famille III, et les biphenyle- et toluène-dioxygénases constituent la famille IV. Une cinquième famille, décrite par Kweon, rassemble les dioxygénases des bactéries Gram⁺ dégradant les HAP et possédant une ferrédoxine [3Fe-4S] (Kweon *et al.*, 2008)

3.3.1 Dioxygénases des bactéries à Gram négatif

Chez les bactéries Gram⁻ les dioxygénases hydroxylant les HAP peuvent être classées en trois sous-groupes suivant leur similarité de séquences : les HAP-dioxygénases, les biphenyle-dioxygénases et les dibenzofurane-dioxygénases (Figure I.21).

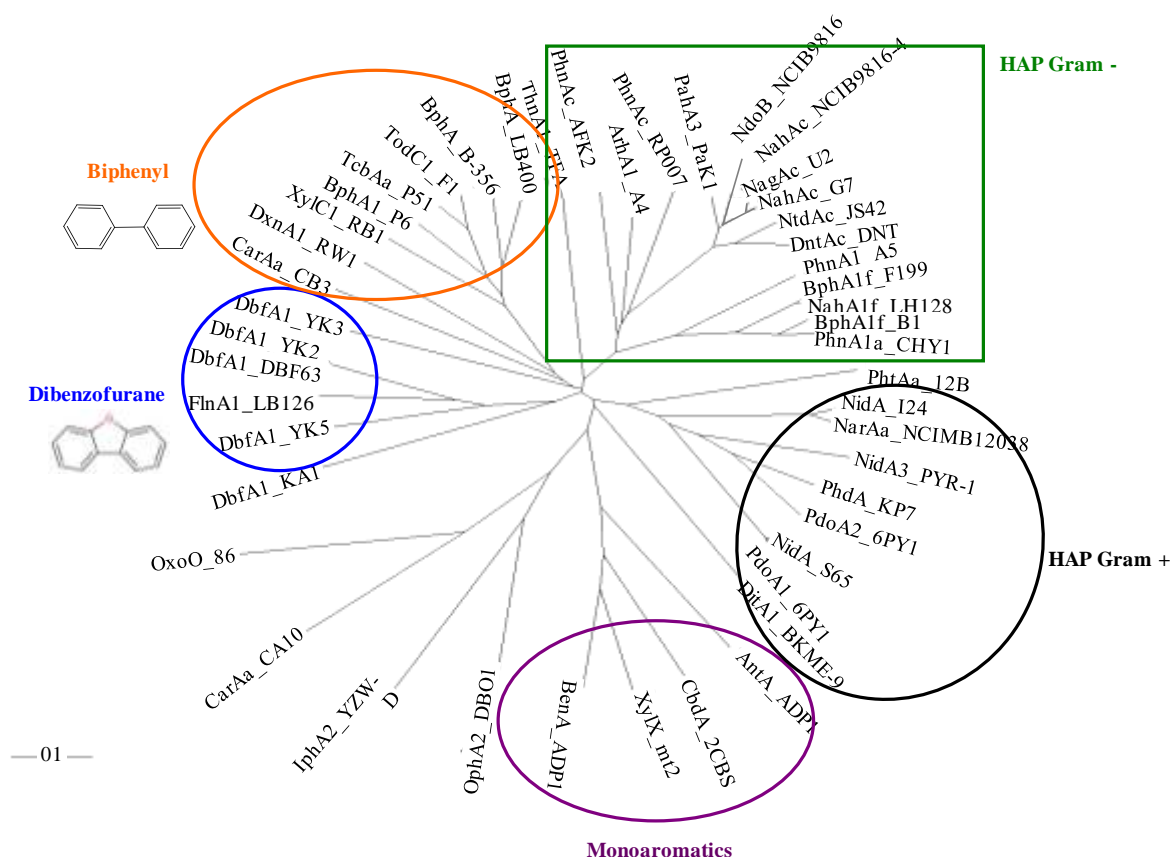


Figure I.21 : Organisation phylogénétique des principales dioxygénases hydroxylant des noyaux aromatiques d'HAP caractérisées chez les bactéries avec leur substrat d'isolement. PhnAc_AFK2, phenanthrene, *Alcaligenes faecalis* AFK2 (AB024945); ArhA1_A4, acenaphthene and acenaphthylene, *Sphingomonas* sp. strain A4 (BAD34447); PhnAc_RP007, phenanthrene, *Burkholderia* sp. RP007 (AF061751); PahA3_PaK1, naphthalene, *Pseudomonas aeruginosa* PaK1 (D84146); NagAc_U2, naphthalene, *Ralstonia* sp. U2 (AF036940); NahAc_G7, naphthalene, *Pseudomonas putida* NCIB9816 (YP_534822); NdoB_NCIB9816, naphthalene, *Pseudomonas putida* NCIB9816 (M23914); NahAc_NCIB9816-4, naphthalene, *Pseudomonas putida* NCIB9816-4 (NP_863072); NtdAc_JS42, 2-nitrotoluene, *Pseudomonas* sp. JS42 (U49504); DntAc_DNT, 2,4-dinitrotoluene, *Burkholderia* sp. DNT (U62430); PhnA1_A5, naphthalene, *Cyclocasticus* sp. strain A5 (AB102786); BphA1f_F199, biphenyl, *Sphingomonas aromaticivorans* F119 (AF079317); NahA1f_LH128, phenanthrene, *Sphingomonas* sp. strain LH128 (EU024112); BphA1f_B1, biphenyl, *Sphingomonas yanoikuyae* B1 (ABM91740); PhnA1a_CHY1, chrysene, *Sphingomonas* sp. strain CHY-1 (AJ633551); PhtAa_12B, phthalate, *Arthrobacter keyseri* 12B (AF331043); NidA_I24, indene, *Rhodococcus* sp. I24 (AF121905); NarAa_NCIMB12038, naphthalene *Rhodococcus* sp. NCIMB12038 (AF082663); NidA3_PYR-1, pyrene, *Mycobacterium vanbaalenii* PYR-1 (AAY85176); PhdA_KP7, phenanthrene, *Nocardioide* sp. KP7 (AB017794); PdoA2_6PY1, pyrene, *Mycobacterium* sp. strain 6PY1 (CAD38643); NidA_PYR-1, pyrene, *Mycobacterium vanbaalenii* PYR-1 (AF249301); NidA_S65, pyrene, *Mycobacterium* sp. S65 (AF 546904); PdoA1_6PY1, pyrene, *Mycobacterium* sp. strain 6PY1 (CAD38647); Dita1_BKME-9, diterprenoid, *Pseudomonas abietaniphila* BKME-9 (AF119621); AntA_ADPI, anthranilate, *Acinetobacter calcoaceticus* ADPI (AF071556); Cbda_2CBS, 2-halobenzoate, *Burkholderia cepacia* 2CBS (X79076); XylX_mt2, toluate, *Pseudomonas*

3.3.1.2 Les Gammaprotéobactéries

Les dioxygénases de *Pseudomonas* ont essentiellement été étudiées à partir d'espèces capables d'utiliser le naphthalène, Nah P.putida G7, Ndo P.putida NCIB 9816-4, Nag P.U2, ou le phénanthrène, Pah P.putida OUS82 (Takizawa *et al.*, 1994). Leur gamme de substrat est restreinte.

La dioxygénase de *Cycloclasticus* sp. A5, PhnA1A2, une Gammaprotéobactérie des milieux marins, a une gamme de substrats hydroxylables similaire aux dioxygénases des Sphingomonades. Elle est capable d'oxyder le biphenyle, le naphthalène, le méthyl-naphthalène, le phénanthrène, le dibenzofurane et le dibenzothiophène. En revanche, elle n'attaque ni l'anthracène, ni le pyrène, ni le benzo[a]pyrène (Kasai *et al.*, 2003).

3.3.1.3 Les Alphaprotéobactéries

Les dioxygénases de Sphingomonades sont capables d'initier l'oxydation d'une large gamme de substrats.

La dioxygénase de *Sphingobium yanoikuyae* B1, BphA1fA2f_B1, dont la sous-unité α présente 90% de similarités de séquence avec celle de *Sphingomonas* CHY-1, utilise préférentiellement le biphenyle, mais peut aussi transformer le naphthalène en 1,2 diol, le phénanthrène en 1,2 et 3,4 diol, le benzo[a]pyrène en 9,10 et 7,8 diols, et le benz[a]anthracène en trois dihydrodiols (Yu *et al.*, 2007).

La dioxygénase de *Sphingomonas* sp. LH128, NahA1fA2f, dont la sous-unité α partage 78% de similarités avec celle de *Sphingomonas* CHY-1 et 98% de similarités avec celle de *Novosphingobium aromaticivorans* sp. F199, BphA1fA2f_F199, présente une plus faible affinité pour les HAP de haut poids moléculaire. Le naphthalène est le substrat préféré de NahA1fA2f, mais la dioxygénase peut aussi oxyder le phénanthrène, le biphenyle, l'anthracène, le fluorène et ses analogues hétéroatomiques, le dibenzofurane, le dibenzothiophène et le carbazole ainsi que le fluoranthène en un diol instable, le chrysène et le benz[a]anthracène en deux dihydrodiols (Schuler *et al.*, 2009). La conservation partielle des acides aminés des deux boucles LI et LII décrites pour *Sphingomonas* CHY-1 pourrait être une des causes de la non oxydation du benzo[a]pyrène par NahA1fA2f. Malgré la large

gamme d'HAP oxydables par cette dioxygénase, *Sphingomonas* sp. LH128 ne peut minéraliser ni le fluorène ni l'anthracène, qu'elle co-métabolise.

Sphingomonas sp. A4 n'est capable de minéraliser que l'acénaphtylène et acenaphtène. Sa dioxygénase, ArhA1 peut, par contre, initier l'attaque du naphthalène, du phénanthrène, de l'anthracène et du fluoranthène (Pinyakong *et al.*, 2004).

Enfin, la dioxygénase de *Sphingomonas* sp. LB126, FlnA1A2, est assez particulière puisqu'elle est proche de celle de *Terrabacter* sp. DBF63, DbfA1A2, une bactérie Gram⁺. Cette dioxygénase angulaire oxyde le fluorène, ainsi que les analogues hétéroatomiques du fluorène. Elle oxyde aussi le phénanthrène en positions 9,10 créant un diol non métabolisable. L'anthracène est hydroxylé en trois produits, un *cis*-1,2-diol, un dérivé monohydroxy et un *cis*-2,3-diol (Schuler *et al.*, 2008).

3.3.1.4 Les Betaprotéobactéries

Les dioxygénases des betaprotéobactéries ont été nettement moins étudiées notamment pour ce qui est de leur spécificité du substrat. Les souches dont elles sont issues ont été caractérisées pour leur capacité à minéraliser certains HAP (voir le paragraphe 2.6).

3.3.2 Dioxygénases des bactéries à Gram positif

Les bactéries Gram⁺ possèdent, quant à elles, plusieurs copies de gènes cataboliques leur conférant une large gamme de spécificité. Ainsi le système décrit chez *Mycobacterium* PYR-1 comporte NidBA, dont le substrat préféré est le pyrène, et NidA3B3 qui oxyde préférentiellement le fluoranthène, mais qui oxyde aussi le naphthalène, le phénanthrène, l'anthracène et le pyrène (Khan *et al.*, 2001; Kim *et al.*, 2006). *Mycobacterium* 6PY1 possède à ce jour Pdo1 qui oxyde préférentiellement le pyrène, ainsi que le phénanthrène, et Pdo2 qui oxyde les HAP de 2 à 3 cycles avec une préférence pour le phénanthrène (Krivobok *et al.*, 2003).

II. Présentation du travail de thèse

Ce projet a pour objectif de caractériser la diversité des communautés bactériennes spécifiquement impliquées dans la biodégradation des hydrocarbures aromatiques polycycliques (HAP) au sein d'un sol qui reçoit, de manière chronique, des eaux de ruissellement d'autoroute chargées en hydrocarbures.

La stratégie mise en œuvre s'appuie sur une méthode de marquage isotopique, le Stable Isotope Probing (SIP), associée à différents outils moléculaires et protéomiques (Figure II.1). Contrairement aux méthodes de microbiologie dépendantes de la culture qui ciblent moins de 5% de la diversité bactérienne, et aux méthodes moléculaires classiques qui ne permettent pas de mettre directement en relation la structure et la fonction des microorganismes, la méthode SIP présente l'avantage d'identifier spécifiquement les communautés bactériennes impliquées dans une fonction métabolique précise.

La stratégie SIP a donc été utilisée afin de répondre à 4 questions :

Dans quelle mesure les HAP sont-ils dégradés dans le sol d'étude ?

Quels sont les acteurs de la biodégradation des HAP ?

Les populations bactériennes impliquées dans la dégradation des HAP évoluent-elles au cours du temps ?

Quelles sont les enzymes cataboliques (dioxygénases) impliquées dans la biodégradation ?

L'approche par marquage isotopique consiste à incuber, en microcosmes, les bactéries du sol en présence d'un HAP modèle, ici le phénanthrène, uniformément marqué avec un isotope stable, le ^{13}C . Au cours de l'incubation, les communautés bactériennes capables de dégrader le substrat, transforment une partie du carbone en CO_2 et incorporent l'autre partie dans leurs constituants cellulaires (ADN, ARN, protéines, lipides).

Le CO_2 marqué dégagé dans les microcosmes a été dosé au cours du temps afin de connaître la cinétique de biodégradation du phénanthrène *in situ*.

L'ADN génomique ^{13}C , plus dense, a été isolé de l'ADN génomique non marqué par ultracentrifugation sur gradient de densité. Les acteurs spécifiques de la biodégradation du HAP modèle ont alors été identifiés *via* la séquence de leur gène codant l'ARN ribosomal 16S, marqueur de l'identité bactérienne. D'autre part, le potentiel catalytique des

communautés bactériennes a également été étudié *via* l'analyse de gènes codant des enzymes métaboliques de type dioxygénases.

En outre, nous avons cherché à mettre en œuvre une méthode d'analyse directe du potentiel métabolique de dégradation des HAP dans le sol, par une approche protéomique.

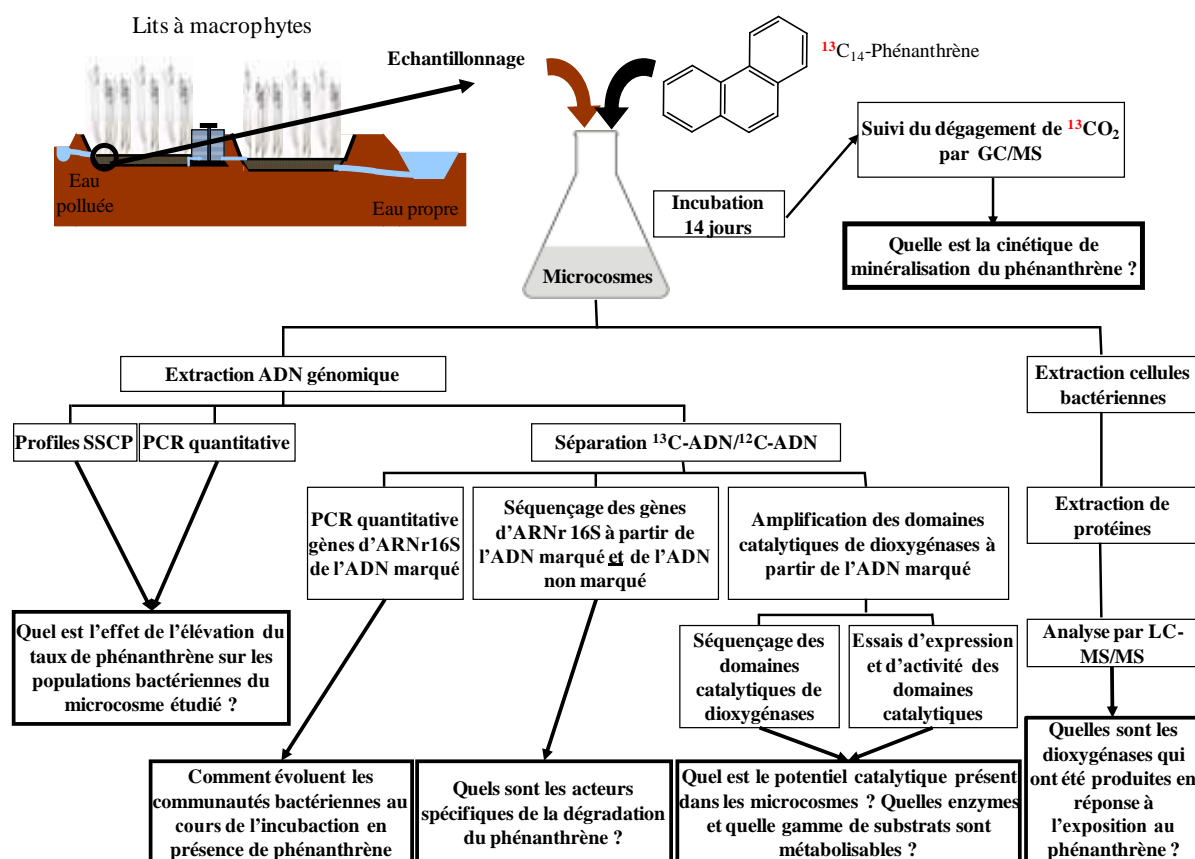


Figure II.1 : Stratégie de marquage isotopique (Stable Isotope Probing) mise en œuvre dans l'exploration de la diversité bactérienne impliquée dans la dégradation des HAP, et l'analyse du potentiel métabolique de ces bactéries.

La présentation des résultats de mes travaux de thèse est divisée en cinq chapitres qui traitent des points suivants :

Le chapitre I présente la synthèse du substrat marqué, la caractérisation du système d'étude et les expériences préliminaires qui ont permis de définir les conditions expérimentales.

Le chapitre II met en évidence la biodégradation du phénanthrène marqué par les bactéries du sol et l'impact sur l'ensemble des populations exposées.

Le chapitre III s'intéresse à l'identification des acteurs de la dégradation du phénanthrène et au suivi de leur devenir en fonction du temps et de la concentration en HAP.

Le chapitre IV présente l'approche moléculaire et biochimique qui a été développée afin d'évaluer le potentiel catalytique des acteurs de la biodégradation des HAP et plus spécifiquement, de caractériser les arène-dioxygénases, responsables de la première étape de dégradation des HAP.

Enfin, le chapitre V est consacré à la mise au point d'une approche protéomique permettant l'analyse et l'identification d'enzymes cataboliques spécifiques, notamment des dioxygénases, directement à partir du sol.

III. Matériel et Méthodes

1. Sols d'étude, échantillonnages, teneur en HAP

1.1 Provenance des échantillons

Les échantillons de sol utilisés dans cette étude proviennent d'un système de traitement des eaux pluviales recueillies après qu'elles se sont chargées des polluants (hydrocarbures) qui se déposent sur les chaussées d'autoroute. L'installation, située au bord de l'A43 Nord près de Chambéry (Savoie), est composée de deux bassins successifs, le deuxième se déversant dans un ruisseau local. Les eaux de chaussée sont acheminées jusqu'au premier bassin de rétention dont le fond est recouvert d'une couche argileuse imperméable (Figure III.1-A). La structure du second bassin de type lit à macrophytes comporte un système filtrant à base de graviers sur géotextile (Figure III.1-B) sur lequel des roseaux ont été plantés (*Phragmites australis*). D'autres plantes aquatiques (*Typha latifolia* et *Phragmites australis*) ont colonisé naturellement le premier bassin. A la base des bassins, un système de drains sur lit d'argile les isole du sol sous-jacent et permet l'écoulement vers le ruisseau.

Les échantillons de sols utilisés pour les microcosmes proviennent du bassin de rétention 1 en amont de la vanne de communication avec le second bassin, équipée d'un séparateur d'huile. Ils ont été prélevés en quatre points (500 g/prélèvement), mélangés afin d'obtenir un échantillon représentatif homogène (Figure III.1) et placés dans des boîtes plastiques et dans une glacière pour le transport. Les échantillons proviennent de la couche supérieure de 10 cm saturée d'eau et ont l'aspect d'une boue riche en matière organique d'origine végétale. Les boues ont été aliquotées dans des tubes Falcon 50 ml (10 g/tube) et stocké à 4°C avant utilisation.

1.2 Extraction des HAP

Afin de connaître la teneur en HAP du sol d'étude, des échantillons d'environ 1 g humide ont été préparés dans des tubes en verre de 18 ml. Une masse équivalente de sulfate de sodium anhydre a été ajoutée à chaque échantillon, ainsi que 10 ml d'hexane. Le mélange a été homogénéisé au vortex puis placé dans un bain à ultrasons pendant 1 h. Il a ensuite été placé sur un agitateur horizontal à l'abri de la lumière, durant 16 h. Les phases organiques ont été transférées dans des tubes Falcon de 15 ml. Après centrifugation et transfert dans un nouveau

tube, le résidu a été extrait une seconde fois avec 10 ml d'hexane. Le solvant a été évaporé sous argon et l'extrait sec a été repris dans 0,5 ml d'acétonitrile. Les HAP ont été dosés par GC/MS.

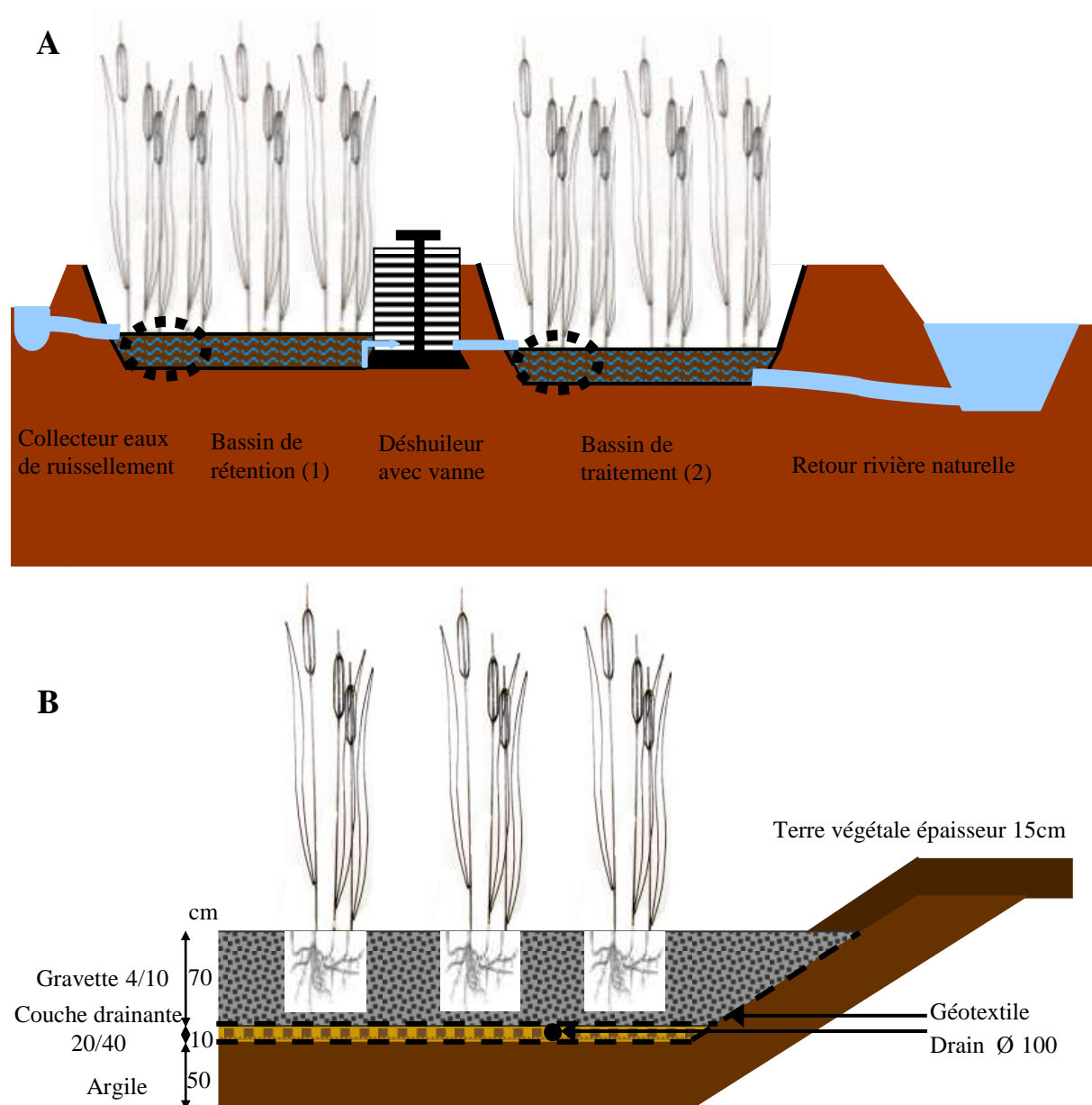


Figure III.1 : Schéma de principe des bassins de rétentions de l'AREA sur l'autoroute A43 Nord près de Chambéry (Savoie). A : Les cercles en pointillé noir représentent les zones de prélèvement. B : Le schéma en coupe montre la structure du second bassin de type lit à macrophytes.

2. Mise en place des microcosmes pour les expériences de marquage *in situ* par un isotope stable

2.1 Essais préliminaires

Des microcosmes ont été préparés dans des flacons Erlenmeyer de 250 ml stériles avec 5 g de sol dans chaque flacon. Une solution de phénanthrène 1 mg/ml (5,6 mM) dans l'acétone a été ajoutée à raison de 160 µg ou 1 mg de phénanthrène par flacon. Ces essais, référencés Ex50 et Ex300, ont été réalisés en duplicats (A et B). Après évaporation du solvant sous une Sorbonne, 45 g de sol ont été rajoutés et mélangés dans chaque flacon, ainsi que 10 ml d'H₂O stérile. Un microcosme de contrôle contenant 50 g de sol et un microcosme abiotique contenant 50 g de sol et 0,2% d'azoture de sodium ont aussi été préparés. Les flacons ont été hermétiquement fermés à l'aide de bouchons en caoutchouc, et incubés 22 jours à 25°C sans agitation et dans le noir. Des échantillons de sol ont été prélevés avant (T₀) et à la fin de l'incubation (T_f).

2.2 Microcosmes pour la série d'expérience SIP1

Tous les microcosmes contenaient 20 g de sol humide auxquels ont été ajoutés 5 ml de solution saline et du phénanthrène comme indiqué dans le Tableau III.1. Ils ont été préparés dans des flacons Erlenmeyer de 250 ml stériles.

Le phénanthrène non marqué (¹²C) ou bien intégralement marqué (¹³C) a été ajouté en solution dans le DMSO (10 mg/ml).

La solution saline (7,5 mM (NH₄)₂SO₄ – 20 mM KH₂PO₄ – 30,6 mM Na₂HPO₄ – 0,18 mM CaCl₂ – 3,6 10⁻² mM FeSO₄ – 0,81 mM MgSO₄) a été ajoutée pour palier une éventuelle carence en sels minéraux du sol.

Enfin, la souche *Sphingomonas* sp. CHY-1 a été inoculée dans deux microcosmes à raison d'environ 2 x 10⁸ bactéries/flacon. Les bactéries ont été cultivées sur milieu MSM phénanthrène puis lavées en milieu salin pour éliminer le phénanthrène résiduel, avant d'être mises en suspension à une DO₆₀₀ = 1 pour inoculation.

Les flacons ont été hermétiquement fermés par des bouchons en caoutchouc, puis incubés à 25°C sans agitation à l'obscurité.

Tableau III.1 : Composition des microcosmes – expériences SIP1

Microcosme	Réplicas	Ajouts	Teneur Phénanthrène
Abiotique	2	120 µl d'azoture de sodium 5 M (0,2% final)	Concentration du sol
Contrôle négatif	2	Pas d'ajout	Concentration du sol
Contrôle positif	2	25 µl de ¹² C-phénanthrène	10 µg/g de sol en plus de la concentration initiale
SIP200	3	27 µl de ¹³ C-phénanthrène	13,4 µg/g de sol en plus de la concentration initiale
SIP2000	3	270 µl de ¹³ C-phénanthrène	134 µg/g de sol en plus de la concentration initiale
SIP2000 CHY1	2	270 µl de ¹³ C-phénanthrène 0,2 ml <i>Sphingomonas</i> CHY-1 (2 x 10 ⁸ bactéries)	134 µg/g de sol en plus de la concentration initiale

3. Techniques d'analyses physico chimiques

3.1 Dosage des HAP et des dérivés hydroxylés par chromatographie gazeuse et spectrométrie de masse

L'analyse des HAP et de leurs produits de dégradation a été réalisée par chromatographie en phase gazeuse (GC) à l'aide d'un appareil HP6890 (Agilent Technologies) équipé d'une colonne MDN12 (30 m x 0,25 mm, 25 µm d'épaisseur de phase stationnaire ; Supelco) dans laquelle passe de l'hélium ultra pur, le gaz vecteur, à 1 ml/min. L'appareil GC est couplé à un détecteur de masse HP 5973 MS (Agilent Technologies) fonctionnant à une énergie d'ionisation de 70eV et une tension de 2 000 volts. L'appareil a été réglé en mode Single-Ion Monitoring (SIM), pour augmenter la sensibilité de détection des ions correspondant aux masses des produits cibles à analyser (m/z théorique).

3.1.1 Dosage des HAP

Les échantillons (2,5 µl) ont été injectés automatiquement dans un dispositif où ils ont été volatilisés. Une fraction du volume injecté (1/50) est ensuite entrée dans la colonne avec un débit de 50 ml/min. La température du four a été augmentée de 75 à 300°C selon un gradient

de 17°C/min de pas, puis maintenue 10 min à 300°C. Ce programme a été optimisé pour la séparation des 16 HAP du mélange standard (Krivobok *et al.*, 2003). Les HAP ont été élués entre 8,9 et 32,7 min et détectés par le spectromètre de masse afin d'être identifiés.

La concentration des HAP a été calculée sur la base d'une courbe de calibration obtenue à partir d'un mélange standard des 16 HAP (fourni par Agilent).

3.1.2 Dosage des dérivés hydroxylés

Les dérivés hydroxylés des HAP ont été dérivatisés selon une méthode décrite dans le paragraphe 7.3 de Matériels et Méthodes. Ils ont ensuite été injectés automatiquement (2,5 µl) dans le dispositif de volatilisation puis une fraction (1/5) du volume est entré dans la colonne à un débit de 5 ml/min. La température du four a été augmentée de 75 à 300°C selon un gradient de 17°C/min de pas, puis maintenue 7,76 min à 300°C. Les dérivés hydroxylés ont été élués entre 9,25 et 13,1 min et détectés par le spectromètre de masse afin d'être identifiés.

La concentration des dérivés hydroxylés a été calculée sur la base d'une courbe de calibration obtenue à partir des dérivés diols purifiés au laboratoire (Jouanneau and Meyer, 2006).

3.2 Dosage du CO₂ par chromatographie gazeuse et spectrométrie de masse

La minéralisation de la matière organique et plus spécifiquement du phénanthrène dans les microcosmes a été suivie par dosage du CO₂ dégagé dans les flacons hermétiquement fermés. Cette mesure a été réalisée par GC/MS avec le même appareil que ci-dessus équipé d'une colonne HP-PLOT/Q (15 m x 0,32 mm, 20 µm d'épaisseur de phase ; Agilent) maintenue sous flux d'hélium à 36 ml/min. Après injection manuelle de l'échantillon (25 µl), une fraction (1/12) est entrée dans la colonne. La température du four a été maintenue à 40°C. Le détecteur de masse HP 5973 MS a été réglée en mode Single-Ion Monitoring (SIM) à une énergie d'ionisation de 70eV et une tension de 1 400 volts. Les deux types de CO₂ ont été élués après 1 min. L'utilisation du mode SIM a donc permis de détecter séparément les masses du ¹²CO₂ (m/z = 44) et du ¹³CO₂ (m/z = 45) lors de deux mesures successives.

Un mélange Argon : CO₂ (90% : 10%) a été dilué dans des flacons remplis d'argon afin de construire une gamme de calibration. L'aire sous chaque pic correspondant au ¹²CO₂ (m/z = 44) a été reporté en fonction de la concentration connue du mélange.

A chaque point de la cinétique de minéralisation, les aires de pics du ¹²CO₂ et du ¹³CO₂ mesurées dans l'air ambiant ont été soustraites à celles du ¹²CO₂ et du ¹³CO₂ mesurées dans les microcosmes. Les aires de pic obtenues ont été converties, à partir de la courbe de calibration, en concentration de CO₂.

3.3 Dosage du phénanthrène en solution dans le DMSO

Lors de la mise en place des microcosmes, le phénanthrène non marqué et le phénanthrène marqué ont été dissouts dans une solution de DMSO afin d'obtenir une concentration de 10 mg/ml. La concentration des solutions stock a été contrôlée par mesure de l'absorbance UV (200-350 nm) à l'aide d'un Diode Array Spectrophotometer Hewlett Packard 8452A. La concentration a été calculée à partir des absorbances à 250 nm, 280 nm et 291 nm, avec les coefficients d'extinction molaire connus dans le méthanol (David, 1992)

$$\epsilon_{250} = 64\,100 \text{ cm}^{-1} \cdot \text{M}^{-1}$$

$$\epsilon_{280} = 8\,550 \text{ cm}^{-1} \cdot \text{M}^{-1}$$

$$\epsilon_{291} = 12\,250 \text{ cm}^{-1} \cdot \text{M}^{-1}$$

Les solutions stock ont été diluées au 1/1600^e et au 1/3200^e (d). L'absorbance (A) aux différentes longueurs d'onde (λ) a été mesurée dans une cuve de 500 μ l et de chemin optique (l) 0,2 cm. La concentration est calculée par la loi de Beer-Lambert :

$$C = d * \frac{A_{\lambda}}{l * \epsilon_{\lambda}}$$

3.4 Analyse LC/MS-MS des protéines du sol

Les analyses des protéines extraites du sol ont été réalisées à la plateforme de protéomique du Laboratoire d'Etude de la Dynamique des Protéomes (CEA/iRTSV/Laboratoire Biologie à Grande Echelle, Grenoble)

3.4.1 Digestion des protéines

Les bandes de gel contenant les protéines d'intérêt ont été excisées manuellement et nettoyées par plusieurs incubations de 15 min dans 25 mM NH_4HCO_3 puis 15 min dans 25 mM NH_4HCO_3 contenant 50% (v/v) acétonitrile. Les bouts de gel ont alors été déshydratés dans l'acétonitrile (100 %) et incubés avec H_2O_2 (7%) pendant 15 min avant d'être lavés à nouveau avec les solutions décrites au début. 0,15 μg de trypsine modifiée (Promega, sequencing grade) reprise dans 25 mM NH_4HCO_3 a été ajouté aux bouts de gel déshydratés et incubés toute la nuit à 37°C. Les peptides ont alors été extraits des bouts de gel en 3 étapes successives dans 30 μl d'acétonitrile (50%), 30 μl d'acide formique (5%) et 30 μl d'acétonitrile (100%). Les surnageants ont alors été rassemblés et séchés sous vide.

3.4.2 Analyses en chromatographie liquide couplé à un tandem de spectrométrie de masse (nano-LC-MS/MS)

Les peptides extraits et séchés ont été repris dans l'acétonitrile (5%) et l'acide trifluoroacétique (0,1%) puis analysés par nano-LC-MS/MS en ligne (Ultimate 3000, Dionex et LTQ-Orbitrap, Thermo Fischer Scientific). La méthode nano-LC a consisté en un gradient de 40 min de 5% à 40% d'acétonitrile dans 0,1% acide formique à un débit de 300 nl/min. Les peptides ont été chargés sur une pré-colonne C18 PepMap de 300 μm x 5 mm et séparés sur une colonne C18 de 75 μm x 150 mm (Gemini C18, Phenomenex). Les données MS et MS/MS ont été acquises en utilisant Xcalibur (Thermo Fischer Scientific) et traitées automatiquement en utilisant le programme Mascot Daemon (version 2.2, Matrix Science).

3.4.3 Analyses bioinformatiques

Des recherches contre la banque SwissProt-Trembl_decoy database (taxonomie Bacteria) ont été réalisées pour chaque échantillon en utilisant une version locale de Mascot 2.2. ESI-TRAP a été choisi comme instrument, trypsin/P comme enzyme et deux mauvais-clivages permis. Les tolérances d'erreur de masse pour les précurseurs et les fragments ont été respectivement 10 ppm and 1 Da. Les modifications peptidiques variables autorisées dans les requêtes ont été : acetyl (N-ter), dioxidation (M), oxidation (M) et trioxidation (C). Le programme IRMa (Bruley *et al.*, 2009) a été utilisé pour filtrer les résultats : query homology threshold $p < 0,01$ et minimum de deux peptides différents pour considérer une protéine comme identifiée.

4. Techniques de microbiologie

4.1 Souches bactériennes utilisées

Sphingomonas sp. CHY-1 et *Mycobacterium* sp. 6PY1 ont été isolées au laboratoire à partir de sol pollués provenant d'un site où était installée une ancienne usine à gaz (Tableau III.2). La souche CHY-1 est capable d'utiliser le chrysène comme seule source de carbone, et peut croître sur naphtalène, phénanthrène ou anthracène (Willison, 2004). Plusieurs enzymes de dégradation des HAP, notamment la naphtalène-dioxygénase, ont été étudiées dans cette souche (Demaneche *et al.*, 2004; Jouanneau *et al.*, 2006; Jouanneau *et al.*, 2007). La souche 6PY1 peut utiliser le pyrène et le phénanthrène comme substrats de croissance et dégrade le fluoranthène. Certaines enzymes impliquées dans la dégradation ont été identifiées (Krivobok *et al.*, 2003). La souche de *Acidovorax* sp. NA3, récemment isolée pour son aptitude à dégrader le phénanthrène, a été fournie par D.R. Singleton (Singleton *et al.*, 2009). Plusieurs souches de *Escherichia coli* ont été employées selon les méthodes de clonage et d'expression mises en œuvre (Tableau III.2).

Tableau III.2 : souches bactériennes utilisées dans cette étude

Souches	Génotype/Phénotype	Utilisation	Référence Fournisseur
<i>Escherichia coli</i> DH5α	<i>F</i> – Φ80 <i>lacZ</i> Δ <i>M15</i> Δ(<i>lacZ</i> YA- <i>argF</i>) <i>U169</i> <i>recA1 endA1 hsdR17</i> (<i>rK</i> –, <i>mK</i> +) <i>phoA supE44 λ</i> – <i>thi</i> – <i>1 gyrA96 relA1</i>	Clonage/ Surexpression	Gibco
<i>Escherichia coli</i> Top10	<i>F</i> – <i>mcrA</i> Δ(<i>mrr</i> - <i>hsdRMS</i> - <i>mcrBC</i>) Φ80 <i>lacZ</i> Δ <i>M15</i> Δ <i>lacX74 recA1 araD139</i> Δ(<i>ara leu</i>) 7697 <i>galU galK</i> <i>rpsL</i> (<i>StrR</i>) <i>endA1 nupG</i>	Clonage	Invitrogen
<i>Escherichia coli</i> JM109	<i>recA1, endA1, gyrA96, thi</i> , <i>hsdR17, supE44, relA1</i> , Δ(<i>lac-proAB</i>)/ <i>F'</i> [<i>traD36</i> , <i>proAB</i> ⁺ , <i>lacI</i> ^q , <i>lacZ</i> Δ <i>M15</i>]	Surexpression	(Yanisch-Perron <i>et al.</i> , 1985)
<i>Escherichia coli</i> NEB5-alpha (Subcloning efficiency)	<i>fhuA2</i> Δ(<i>argF-lacZ</i>) <i>U169</i> <i>phoA glnV44</i> Φ80Δ (<i>lacZ</i>) <i>M15 gyrA96 recA1</i> <i>relA1 endA1 thi-1 hsdR17</i>	Clonage	New England BioLabs
<i>Escherichia coli</i> NEB5-alpha (High efficiency)	<i>fhuA2</i> Δ(<i>argF-lacZ</i>) <i>U169</i> <i>phoA glnV44</i> Φ80 Δ(<i>lacZ</i>) <i>M15 gyrA96 recA1</i> <i>relA1 endA1 thi-1 hsdR17</i>	Clonage	New England BioLabs
<i>Escherichia coli</i> NovaBlue Singles™	<i>endA1 hsdR17</i> (<i>rK12</i> [–] <i>mK12</i> ⁺) <i>supE44 thi-1 recA1</i> <i>gyrA96 relA1 lac</i> <i>F'</i> [<i>proA</i> ⁺ <i>B</i> ⁺ <i>lacI</i> ^q Δ <i>M15::Tn10</i>] (<i>Tet</i> ^R)	Clonage	Novagen
<i>Acidovorax</i> sp.NA3		Souche modèle dégradant les HAP	(Singleton <i>et al.</i> , 2009)
<i>Sphingomonas</i> sp. CHY-1		Souche modèle dégradant les HAP	(Demaneche <i>et al.</i> , 2004; Willison, 2004)
<i>Mycobacterium</i> sp.6PY1		Souche modèle dégradant les HAP	(Krivobok <i>et al.</i> , 2003)
<i>Pseudomonas</i> sp. KT2442		Souche modèle dégradant les HAP	(Franklin <i>et al.</i> , 1981)

4.2 Milieux de culture

4.2.1 E.coli et Pseudomonas

4.2.1.1 Milieu LB

Tryptone 10 g/l – extrait de levure 5 g/l – NaCl 10 g/l – ajuster le pH à 7.5 avec une solution NaOH 10 M

Stériliser à l'autoclave : 121°C, 15 min.

Pour la préparation de milieux solides, ajouter 15 g/L d'agar avant stérilisation à l'autoclave

Les incubations sont réalisées à 37°C sous agitation 200 rpm.

4.2.1.2 Milieu M9

Milieu M9 concentré (5 x)

NH₄Cl 5,0 g/l (93mM) – Na₂HPO₄ . 7H₂O 64 g/l (0,24 M) ou Na₂HPO₄ . 12 H₂O 85,5 g/l (0,23 M) – KH₂PO₄ 15 g/l (0,11 M) – NaCl 2,5 g/l (43 mM)

Stériliser à l'autoclave : 121°C, 15 min.

Après stérilisation, diluer avec 4 volumes d'eau stérile et ajouter 1 ml/l de MgSO₄.7H₂O(1M) (stérilisé à l'autoclave : 121°C, 15 min) et 10 ml/l de Glucose 20% (stérilisé par filtration) pour obtenir le milieu M9 complet.

Les incubations sont réalisées à 37°C sous agitation 200 rpm.

4.2.2 Sphingomonas et Mycobacterium

4.2.2.1 Milieu MSM

NH₄Cl 1,2 g/l (22 mM) – K₂HPO₄ 1,6 g/l (9,2 mM) – KH₂PO₄ 0,4 g/l (2,9 mM) – NaCl 0,1 g/l (1,7 mM). Ajuster à pH 7,1 avec HCl et stériliser à l'autoclave : 121°C, 15 min.

Pour la préparation de milieux solides, ajouter 15 g/l d'agar avant stérilisation à l'autoclave

Après stérilisation, ajouter 10 ml/l de chacune des solutions stériles suivantes :

MgSO₄.7H₂O 20 g/l – CaCl₂. 2H₂O 10 g/l – Fe-EDTA (stérilisé par filtration) 1 g/l, extrait de levure Difco 2 g/l. Le substrat de culture est ajouté à cette étape : Na-Benzate (5 mM final) ou succinate (10 mM final) ou Tween80 (2,5 g/l final) ou huile de paraffine (2,5 g/l final). Le

phénanthrène est ajouté en solution dans l'huile de silicone (2 mM final), et dans ce cas, les bactéries se développent en culture biphase. Les incubations sont réalisées à 25°C sous agitation à 200 rpm.

4.2.2.2 Milieu PTYG

Glucose 10 g/l – extrait de levure 10 g/l – peptone 5 g/l – tryptone 5 g/l – $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ 7,5 % 1 ml – $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ 20 % 3 ml. Stériliser à l'autoclave 121°C, 15 min.

Pour les cultures en milieux liquide, les milieux PTYG2 (dilué au demi dans l'eau stérile) ou PTYG20 (dilué au 1/20^e dans l'eau stérile) sont utilisés. Ils sont incubés à 25°C sous agitation à 200 rpm.

4.2.2.3 Milieu PTYG 20 solide

Stériliser 950 ml d'eau contenant 15 g agar. Après refroidissement partiel, ajouter stérilement 50 ml de PTYG. Répartir à raison d'environ 20 ml par boîte de Petri. L'incubation est réalisée à 25°C.

4.2.3 Acidovorax

Sur les recommandations de D.R. Singleton, *Acidovorax* sp. NA3 est cultivé sur un milieu Nutrient Broth liquide ou solide (Difco). Ce milieu contient 3 g/l d'extrait de bœuf et 5 g/l de bactopeptone. Le milieu est stérilisé 15 min à 121°C. Pour le milieu solide, 15 g/l d'agar sont ajoutés au milieu avant stérilisation. La culture est réalisée à 30°C sous agitation à 200 rpm.

5. Biologie moléculaire

5.1 Plasmides utilisés dans cette étude

Les techniques de biologie moléculaire développées lors de cette étude ont utilisé différents plasmides (Tableau III.3, Tableau III.4 et Tableau III.5). Leur utilisation et leur construction seront expliquées plus tard dans ce chapitre ou dans les chapitres de résultats.

Tableau III.3 : Plasmides commerciaux ou disponibles au laboratoire

Plasmide	Description	Résistance	Promoteur et origine de réplication	Référence ou fournisseur
pJET1.2 Blunt	clonage	Ampicilline ¹⁰⁰	Promoteurs T7 - pMB1 origin	Fermentas
pSTBlue-1	clonage	Ampicilline ¹⁰⁰ / Kanamycine ²⁰	Promoteurs T7 et SP6 – f1 origin	Novagen
pUC18	Clonage expression	Ampicilline ¹⁰⁰	Promoteur lac - Origine de réplication pMB1 (incompatible avec ColE1)	Fermentas
pET15b	expression	Ampicilline ¹⁰⁰	Promoteur T7 – ColE1 origin	Novagen
pET9a	expression	Kanamycine ²⁰	Promoteur T7 – ColE1 origin	Novagen
pIZ1036	expression	Kanamycine ²⁰	Promoteur tac – pBBR1CM origin	(Moreno-Ruiz <i>et al.</i> , 2003)
pVLT31	expression	Tetracycline ¹⁰	Promoteur tac – pMMB207 origin	(De Lorenzo <i>et al.</i> , 1993)

Tableau III.4 : plasmides construits au laboratoire lors d'études antérieures

Plasmide	Plasmide d'origine	Gène introduit	Taille du gène introduit (pb)	Référence
pSD8	pET15b	<i>phnA1A2</i>	1931	(Demaneche <i>et al.</i> , 2004)
pSD9	pVLT31	<i>phnA1A2</i>	1931	(Demaneche <i>et al.</i> , 2004)
pIZpod2	pIZ1036	<i>pdoA2B2</i>	1984	Meyer non publié
pEB431	pET15b	<i>phnA4A3</i>	1591	(Demaneche <i>et al.</i> , 2004)
pIBA34	pIZ1036	<i>phnA3A4</i>	1591	Jouanneau, non publié

Tableau III.5 : Plasmides construits pour cette étude

Plasmide	Plasmide d'origine	Insertion	Taille de l'insert (pb)
pFMA1	pUC18	<i>phnA1A2</i>	1931
pFMA2	pFMA1	dcatdiox U3-116*	701
pFMA3	pFMA1	dcatdiox U3-60*	701
pFMA4	pFMA1	dcatdiox U3-72*	701
pFMA5	pFMA1	dcatdiox U3-89*	695
pFMA6	pFMA1	dcatdiox U3-124*	568
pFMA7	pFMA1	dcatdiox U3-16*	568
pFMA8	pFMA1	dcatdiox U3-111*	938
pFMA9	pFMA1	dcatdiox BB10*	944
pFMA10	pFMA1	dcatdiox U3-111*	624
pFMA11	pFMA1	dcatdiox BB10*	624
pLMX3	pSD8	Promoteur tac à la place du T7	31
pLMX4	pLMX3	<i>phnA4A3</i>	1591
pFMA12	pJET1.2	<i>phnAcAd</i>	1916
pFMA13	pET15b	<i>phnAcAd</i>	1916
pFVL1	pUC18	<i>phnAcAd</i>	2072

* : fusion en phase entre les domaines catalytiques amplifiés à partir de l'ADN ¹³C du sol et le domaine N-term de PhnI. Dcatdiox indique le domaine catalytique de la dioxygénase et le nom qui suit fait référence au nom du clone issu des bibliothèques des domaines catalytiques.

5.2 Préparation ADN génomique

5.2.1 *Escherichia coli*

L'ADNg de *E.coli* a été préparé selon une méthode décrite précédemment (Ausubel *et al.*, 1999). Une colonie isolée de *E.coli* DH5 α (sur LB agar) a été inoculée dans 3 ml de LB, 16 h à 37°C. Cette préculture a servi d'inoculum à une culture de 400 ml de LB, incubée 16 h à 37°C. En fin de croissance, la DO₆₀₀ a atteint 4. La culture a été centrifugée 15 min à 10 000 xg puis le culot a été repris dans 38 ml de TE (10 mM Tris HCl pH 8 – 1 mM EDTA), contenant 20 mg de RNase (100 mg/ml), 0,5% de SDS et 4 μg de protéinase K (20 mg/ml). La suspension a été incubée 1 heure à 37°C. Le mélange a ensuite été incubé 20 min à 65°C en présence de 0,6 mM NaCl et 1%/0,07 M de CTAB/NaCl*. La phase aqueuse a été extraite avec une solution de chloroforme/alcool isoamylique (24 : 1) (1 x 1 volume équivalent), puis l'ADN a été précipité en présence de 0,6 volume équivalent d'isopropanol. Il a été lavé avec 4 ml d'éthanol 70%, puis précipité une nouvelle fois par centrifugation 5 min à 10 400 xg . Le culot a été séché à l'air puis repris dans 4 ml de tampon TE (10 mM Tris HCl pH 8 – 1 mM EDTA) (l'agitation à 37°C peut permettre une meilleure dissolution).

L'extraction à partir de 1 ml de culture a été réalisée à l'aide du kit Wizard Genomic® DNA purification kit (Promega)

* Pour préparer la solution stock de CTAB/NaCl, dissoudre 4,1 g de NaCl dans 80 ml d'H₂O, agiter lentement et chauffer. Ajouter 10 g de CTAB petit à petit. Ajuster le volume à 100 ml pour obtenir CTAB/NaCl (10%/0,7 M).

5.2.2 *Sphingomonas*

L'ADN génomique d'une souche pure de *Sphingomonas* a été extrait par une méthode utilisée précédemment au laboratoire (Demaneche *et al.*, 2004).

Une colonie isolée de *Sphingomonas* sp. CHY1 (PTYG20 agar) a été inoculée dans 3 ml de PTYG2 et incubée à 25°C. La préculture a ensuite été transférée dans 200 ml de PTYG2 et incubée à 25°C. Lorsque la DO₆₀₀ a atteint 0,7, la suspension a été centrifugée 15 min à 10 000 xg . Le culot a été repris dans 3,6 ml de tampon TE (10 mM Tris HCl pH 8 – 1 mM EDTA) avec 1,6 mg de RNase (100 mg/ml) et 2 mg de protéinase K (20 mg/mL). Les parois cellulaires ont été fragilisées en présence de lysozyme (2 mg/ml final, solution stock

100 mg/ml) pendant 1 heure à 37°C en mélangeant par inversions. La suspension a été incubée 1 h à 65°C en présence de 1% de SDS (solution stock SDS 10%), 0,5 M de NaCl (solution stock 5 M) et 1% / 0,07 M de CTAB/NaCl. La phase aqueuse a successivement été extraite par un volume de chloroforme/alcool isoamylique (24 : 1), un volume de phénol/chloroforme/alcool isoamylique (25 : 24 : 1) et de nouveau par un volume de chloroforme/alcool isoamylique (24 : 1) : à chaque étape la phase aqueuse a été mélangée puis centrifugée 20 min à 2 057 *xg* et enfin décantée. Elle a ensuite été reprise dans un nouveau tube avec 1 volume d'éthanol 100%. L'ADN a été précipité 1 h à -80°C, centrifugé 30 min à 4°C, 4 626 *xg*, lavé (10 ml d'éthanol 70%) et enfin séché à l'air puis repris dans 100 µl de TE (10 mM Tris HCl pH 8 – 1 mM EDTA).

5.2.3 Mycobacterium

L'ADN génomique d'une souche pure de mycobactéries est extrait par une méthode spécifique utilisée au laboratoire (Gonzalez-y-Merchand *et al.*, 1996).

Une colonie isolée de *Mycobacterium* sp. 6PY1 (PTYG20 agar) a été inoculée dans 3 ml de MSM complet et incubée à 25°C. La préculture a ensuite été transférée dans 200 ml de MSM complet et incubée à 25°C. Lorsque la DO₆₀₀ a atteint 0,7, la suspension a été centrifugée 15 min à 10 000 *xg*. Le culot a été repris dans 10 ml de TEm (Tris HCl 100 mM - pH 8 – EDTA 100 mM – 100 mM NaHPO₄) avec 4 mg de RNase (100 mg/ml) et 2 mg de protéinase K (20 mg/mL). Les parois cellulaires ont été fragilisées en présence de 10 mg/ml lysozyme (solution stock 100 mg/ml) pendant 1 h à 37°C en mélangeant par inversions. La solution a été ajustée à 1% de SDS (1,25 ml de SDS 10%) puis soumise à 3 cycles de congélation (azote liquide 30 s) – décongélation (65°C 10 min). La suspension a été incubée 30 min à 65°C en présence de 0,5 M de NaCl (solution stock 5 M) et de 1%/0,07 M de CTAB/NaCl. La phase aqueuse a été extraite successivement par un volume de chloroforme/alcool isoamylique (24 : 1), 2 volumes de phenol/chloroforme/alcool isoamylique (25 : 24 : 1), et séparée par centrifugation 15 min à 3 214 *xg*. L'ADN a été précipité avec un volume d'isopropanol, centrifugé 20 min à 10°C à 2 057 *xg*, lavé dans 10 ml d'éthanol 70% et repris dans 1 ml de TE (10 mM Tris HCl pH 8 – 1 mM EDTA).

5.2.4 Extraction d'ADN à partir du sol

L'ADN des microorganismes du sol a été isolé à partir d'échantillons de 1 à 10 g de sol humide provenant de microcosmes (voir paragraphe 2 de Matériels et Méthodes). Pour des échantillons de 10 g, l'ADN génomique a été directement extrait du sol à l'aide des kits UltraClean™ Mega Soil DNA ou PowerMax™ Soil DNA Isolation (Mo BIO Laboratories, Ozyme, St Quentin en Yvelines, France). Pour des échantillons de 1 g, le kit PowerSoil™ DNA Isolation (Mo BIO) a été utilisé.

Dans le cas des échantillons de 10 g, le protocole du fabricant a été adapté : pour lyser les bactéries, les échantillons de sol ont été incubés 30 min à 65°C dans des tubes de 50 ml contenant les billes fournies dans le kit, en subissant un brassage mécanique par agitation au vortex 1 min toutes les 10 min. L'extrait soluble débarrassé des principaux constituants du sol a été déposé sur colonne de silice. L'ADN génomique a été élué dans 5 à 8 ml de tampon d'éluion pour obtenir 40 à 50 ng/μl (mesuré à l'aide d'un spectrophotomètre ND-100 NanoDrop Technologies, Inc). Puis il a été précipité en présence d'acétate de sodium (0,3 M final ; solution stock 3 M pH 5,2) et de 2 volumes d'éthanol, et enfin repris dans le tampon TE (10 mM Tris HCl pH 8 – 1 mM EDTA) afin d'obtenir des concentrations d'environ 700 ng/μl.

5.3 Séparation d'ADN génomique marqué ^{13}C sur gradient de densité CsCl

La méthode a été mise au point en s'inspirant de ce qu'ont décrit Lueders *et al.* (Lueders *et al.*, 2004), Martineau *et al.* (Martineau *et al.*, 2008), Manefield *et al.* (Manefield *et al.*, 2002) et Pumphrey *et al.* (Pumphrey and Madsen, 2008b). Des modifications ont été apportées pour optimiser l'efficacité de séparation en utilisant le matériel du laboratoire.

5.3.1 Préparation des gradients

Un tube de centrifugation de contenance 3,3 ml (OptiSeal™ Polyallomer centrifuge tube, Beckman Coulter™) a été rempli avec 2,5 ml de solution de chlorure de césium (solution stock de CsCl : 1,9 g/ml (H₂O) – densité ajustée à 1,82 g/ml), 5 μl de SYBR® Safe (Invitrogen) et 10 à 30 μg d'ADN génomique à séparer. La densité a été calculée par le rapport masse sur volume, en déterminant la masse à 0,1 mg près avec une balance de

précision. Elle a été ajustée à 1,725 g/ml dans un volume final de 3,3 ml en ajoutant soit du tampon (0,1 M Tris HCl pH 8 – 0,1 M KCl – 1 mM EDTA) soit de la solution stock de CsCl. Le gradient a été généré par ultracentrifugation à $413\,000 \times g_{AV}$, 15°C pendant 17 h, en utilisant un rotor TLN-100 et une ultracentrifugeuse OptimaTM TLX (Beckman CoulterTM). Afin de ne pas altérer le gradient, la décélération a été paramétrée pour être la plus douce possible. La séparation de l'ADN lourd et de l'ADN léger a été visualisée sous lampe UV (Mineralight Lamp $\lambda = 254\text{nm}$) et photographiée avec un appareil numérique reflex Sony alpha 100.

5.3.2 Fractionnement du gradient et détection de l'ADN génomique

Le gradient a été fractionné par le bas en perçant le tube avec une aiguille hypodermique et en injectant de l'eau stérile (180 $\mu\text{l}/\text{min}$) par le haut à l'aide d'une pompe péristaltique. Les fractions ont été recueillies dans des tubes de 5 ml à l'aide d'un collecteur Gilson modifié pour maintenir le tube de centrifugation sur la partie mobile du collecteur (Figure III.2). Le contenu du tube a ainsi été séparé en 22 fractions de 150 μl . L'ADN de chaque fraction a été précipité, en présence de glycogène (1 μl de 20 mg/ml, Fermentas), et de polyéthylène glycol/NaCl 30%/1,6 M (300 μl). Après incubation 2 h à température ambiante, la suspension a été centrifugée 1 h à 16 200 xg, le culot a été lavé par 300 μl d'éthanol 70%, séché à l'air et repris dans 20 μl d'H₂O stérile. L'ADN de chaque fraction a été dosé par spectrophotométrie (ND-100), afin d'obtenir une courbe de distribution de l'ADN génomique le long du gradient.

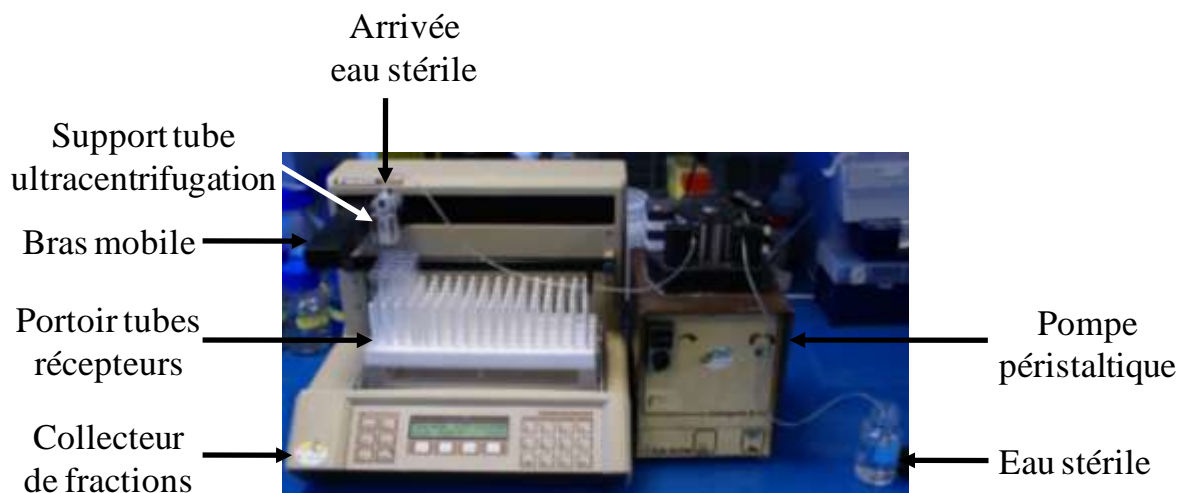


Figure III.2 : Photo du dispositif utilisé pour fractionner le gradient d'ultracentrifugation. Le collecteur de fraction a été modifié par l'ajout de deux pièces en plexiglas : un support élévateur pour mettre le portoir de tubes collecteurs à niveau et un support de tube d'ultracentrifugation fixé au bras mobile. L'eau stérile est acheminée vers le haut du tube d'ultracentrifugation *via* une pompe péristaltique et un tuyau qui se termine par une aiguille hypodermique. Le tube d'ultracentrifugation se remplit goutte à goutte, ce qui a pour effet de pousser le gradient vers le bas. Le tube, percé à son extrémité inférieure, se vide alors goutte à goutte dans les tubes du collecteurs. Toutes les 0,83 min, le bras mobile fait avancer le tube d'ultracentrifugation vers le tube collecteur suivant.

5.4 Amplification, analyse et purification de l'ADN

Les réactions d'amplification d'ADN par PCR (Polymerase Chain Reaction) ont été réalisées avec les amorces et selon les conditions décrites dans le tableau des amorces en Annexe 1. L'enzyme KOD Hot Start DNA Polymerase (Novagen) a été utilisée pour les amplifications requérant une haute fidélité. Le mélange réactionnel se compose de 5 μ l de tampon (10x), 3 μ l de $MgSO_4$ (25 mM), 5 μ l dNTP (2 mM chacun), 1,5 μ l de chacune des amorces (10 μ M chacune), de la matrice d'ADN (100 ng d'ADN génomique ou 2 ng de vecteur) et d'une unité de polymérase dans un volume total de 50 μ l. Après une étape d'activation de l'enzyme, 2 min à 95°C, 25 à 35 cycles d'amplification ont été réalisés, avec une étape de dénaturation, 30 s à 95°C, une étape d'hybridation, 30 s à la température d'hybridation et une étape d'élongation à 70°C variable selon la taille de l'amplicon (programme PCR – Annexe 1).

L'enzyme GoTaq[®] Flexi DNA polymérase (Promega) a été utilisée dans tous les autres cas. Le mélange réactionnel se compose de 5 μ l de tampon Green GoTaq[®] Flexi Buffer (5x) contenant 7,5 mM de $MgCl_2$, 1 μ l de dNTP (10 mM chacun), 1 μ l de chacune des amorces (10 μ M chacune), de la matrice d'ADN (100 ng d'ADN génomique ou 2 ng de vecteur) et de

1,25 unités de polymérase dans un volume total de 25 µl. Après une étape initiale de dénaturation, 2 min à 95 °C, les 25 à 35 cycles d'amplification sont composés d'une étape de dénaturation, 30 s à 1 min à 95°C, d'une étape d'hybridation, 1 min à la température d'hybridation et d'une étape d'élongation à 72°C (programme PCR – Annexe 1). Une élongation finale de 2 à 5 min à 72°C termine l'amplification.

Les produits d'amplification ont été analysés par électrophorèse en gel d'agarose 0,5 à 2 % suivant la taille du fragment, à l'aide d'un dispositif i-Mupid (Eurogentec). Le GelRed™ Nucleic Acid Gel Stain (FluoProbes®) a été incorporé au gel (1/10 000^{ème}) à la préparation. Lors de la migration, il s'intercale entre les bases d'ADN. Les bandes d'ADN ont été révélées par fluorescence à $\lambda = 302$ nm et l'image a été enregistrée grâce à une caméra numérique (Fusion Fx7, Vilber Lourmat). La migration a été réalisée dans le tampon TBE 0,5x (TrisHCl 44,6 mM ; acide borique 44,5 mM ; EDTA 1 mM ; pH8) 1 h à 50 V pour les gels 0,5 et 0,7% ou 30 min à 100 V pour les gels à 1 ou 2%. Les marqueurs de taille GeneRuler™ 100 pb DNA ladder Plus et GeneRuler™ 1 kb DNA ladder (Fermentas) ont été utilisés.

La purification des produits d'amplification ou des fragments de restriction à partir des mélanges réactionnels, ou à partir d'un gel d'agarose, a été réalisée à l'aide du kit NucleoSpin® Extract II (Macherey-Nagel).

5.5 Propagation et purification de l'ADN plasmidique

Les plasmides utilisés lors de cette étude ont été propagés dans des souches de *Escherichia coli* (NEB alpha, Top10 ou DH5α). Selon le gène de résistance porté par le plasmide, les souches transformées ont été cultivées 20 h à 37°C et 200 rpm en milieu LB en présence d'un antibiotique assurant la pression de sélection : Ampicilline (stock 100 mg/ml), Kanamycine (stock 20 mg/ml) ou Tétracycline (stock 10 mg/ml dans 50% éthanol) dilués au 1/1000^e dans le milieu de culture.

Les préparations d'ADN plasmidique ont été réalisées à l'aide des kits NucleoSpin® Plasmid (Macherey-Nagel), pour les MiniPreps, et Nucleobond® (Macherey-Nagel) pour les MidiPreps.

5.6 Techniques de clonage

5.6.1 Préparation des cellules chimio-compétentes

Tampon RF1 : 30 mM acétate de potassium – 100 mM chlorure de potassium – 75 mM chlorure de calcium – 15% (v/v) glycerol bidistillé – ajuster le pH à 5,8 avec une solution d'acide acétique 0,2 M, stériliser par filtration

Tampon RF2 : 0,5 M MOPS pH 6,8 – 10 mM chlorure de potassium – 75 mM chlorure de calcium – 15% (v/v) glycerol bidistillé – ajuster le pH à 6,8 avec une solution d'acide acétique 0,2 M, stériliser par filtration.

A partir d'une boîte de LB-agar inoculée la veille avec la souche d'intérêt et incubée une nuit à 37°C, 5 ml de LB liquide a été inoculé avec une colonie et incubé 2 h à 37°C et 200 rpm. Parallèlement, un Erlenmeyer de 500 ml contenant 50 ml de LB a été préchauffé à 37°C. Au bout de 2 h il a étéensemencé avec la préculture de 5 ml et incubé à 37°C et 200 rpm jusqu'à atteindre une DO_{600} entre 0,4 et 0,8. La culture a ensuite été placée 30 min dans la glace. Pendant ce temps, la centrifugeuse, le rotor et les pots ont été refroidis à 4°C. Les pots utilisés pour la centrifugation doivent avoir un fond large pour faciliter la remise en suspension des cellules bactériennes. La culture a été centrifugée 5 min à 5 000 xg_{AV} et 4°C, et le culot a été lavé dans 20 ml de tampon RF1 glacé. La suspension a été incubée 5 min sur la glace et centrifugée 5 min à 5 000 xg_{AV} et à 4°C. Le culot a été repris dans 2 ml de tampon RF2 glacé, incubé 5 min sur la glace puis aliquoté en fractions de 50 μ l dans des eppendorfs 1,5 ml préalablement refroidis. Les aliquots ont été congelés dans l'azote liquide et conservés à -80°C.

5.6.2 Préparation des cellules électrocompétentes

Une culture de 10 ml de LB a été inoculée avec une colonie issue d'un milieu LB agar et incubée 16 h à 37°C et 200 rpm. Un Erlenmeyer de 2 L contenant 500 ml de LB et préchauffé à 37°C a étéensemencé avec 5 ml de préculture, et incubé à 37°C et 200 rpm jusqu'à atteindre une DO_{600} entre 0,6 et 0,9. La culture a ensuite été placée 30 min dans la glace. Pendant ce temps, la centrifugeuse, le rotor et les pots ont été refroidis à 4°C. La culture a été répartie dans des pots de 250 ml et centrifugée 25 min à 3 300 xg_{AV} et à 4°C. Chacun des

culots a été repris dans 200 ml d'eau stérile glacée. Un second lavage dans 100 ml d'eau stérile glacée a été réalisé. Les culots ont été repris dans 20 ml d'eau glacée stérile contenant 10% de glycérol stérile puis rassemblés dans des tubes Falcon de 50 ml (~40 ml par tube) et centrifugés 10 min à $3\,300\,xg_{AV}$ et à 4°C. Les culots ont été repris dans 1 ml d'eau stérile contenant 10% de glycérol stérile. La suspension concentrée a été aliquotée en fractions de 50 µl dans des eppendorfs 1,5 ml préalablement refroidis, puis congelée dans l'azote liquide et conservée à -80°C.

Pour ces deux protocoles, les centrifugations ont été réalisées dans une Avanti Centrifuge J-30I de Beckman CoulterTM, équipé d'un rotor JLA 9.100, puis dans une centrifugeuse 5804 R (Eppendorf).

Dans chacun des protocoles de clonage suivant, un contrôle négatif a été réalisé afin d'estimer le taux des clones faux positifs. Dans le mélange réactionnel, le volume d'insert a été remplacé par de l'eau.

5.6.3 Clonage classique

Le vecteur de clonage ainsi que le fragment d'ADN à cloner ont été digérés par les enzymes de restriction compatibles (Fermentas), 1 h à 37°C selon les recommandations du fabricant (les enzymes utilisées sont détaillées dans les chapitres de résultats). Si nécessaire, le vecteur a ensuite été traité par la Calf Intestinal Alkaline Phosphatase (CIAP) qui déphosphoryle les extrémités 5', et empêche ainsi le vecteur de se re-circulariser. Les produits de digestion ont été purifiés sur gel.

La ligation entre le vecteur et l'insert a été réalisée par l'enzyme T4 DNA ligase du kit Rapid DNA ligation (Fermentas). En général, un ratio molaire 3 : 1 entre l'insert et le vecteur a été appliqué, dans un mélange de ligation contenant 2 µl de tampon de ligation, 1 µl de T4 DNA ligase et un volume d'eau stérile qsp 10 µl. Après 30 min d'incubation à température ambiante, 50 µl de cellules compétentes (NEB alpha, DH5α ou Top10) ont été transformées par 2 µl du mélange de ligation. Après choc thermique (30 s à 42°C) ou choc électrique (2500 V) les cellules ont été incubées dans 250 µl de LB, 1h à 37°C et étalées sur LB-agar supplémenté avec l'antibiotique dont le gène de résistance est porté par le vecteur.

5.6.4 Clonage de produits de PCR bouts francs

Un kit commercial (Perfectly Blunt® Cloning Kit, Novagen) a été employé pour cloner des fragments d'ADN bouts francs comme ceux générés par l'enzyme KOD. Après vérification de l'amplification sur gel d'agarose, le produit PCR a été purifié par extraction au chloroforme et précipitation à l'isopropanol. La réaction de « End conversion » qui phosphoryle les bout francs de l'insert a été réalisée de manière à obtenir, à l'étape suivante, un ratio molaire minimum de 1 : 2 (vecteur : insert), dans un volume total de 10 µl contenant 5 µl de « End conversion mix ». La réaction a été incubée 15 min à 22°C, puis inactivée 5 min à 75°C et refroidie dans la glace. La ligation dans le plasmide pSTBlue-1 (1 µl soit 50 ng) a été effectuée avec 1 µl d'enzyme T4 DNA ligase (4 unités) 15 min à 22°C. La transformation a été ensuite réalisée en incubant 50 µl de NovaBlue Singles™ Competent Cells (Novagen) avec 2 µl de mélange de ligation 5 min sur la glace, suivi par un choc thermique à 42°C pendant 30s. Les cellules refroidies sur la glace 2 min, ont ensuite été incubées dans 250 µl de milieu de culture (SOC ou LB) 1 h à 37°C, enfin étalées sur LB-Agar + ampicilline. Les boîtes de Pétri ont été incubées à 37°C pendant 16 h.

Dans certains cas, le clonage de fragments bouts francs a été réalisé avec le kit CloneJET™ PCR Cloning Kit (Fermentas). La réaction de ligation comprend le produit PCR (0,15 pmol), purifiés comme décrit précédemment, le vecteur pJET1.2/blunt cloning (0,05 pmol – 50 ng), 10 µl de tampon de réaction (2x), et 1 µl d'enzyme T4 DNA ligase. Le volume réactionnel a été ajusté à 20 µl avec de l'eau (nuclease-free), puis incubé 5 min à la température ambiante. La transformation de 2 µl de mélange de ligation a été réalisée dans les cellules électrocompétentes Top10 par choc électrique à l'aide d'un electoporator 251 (Eppendorf) réglé à 2500 V. Les cellules ont été reprises dans 250 µl de LB, incubée 1 h à 37°C et enfin étalées sur un milieu LB-agar supplémenté avec de l'ampicilline.

Le pJET1.2/blunt cloning possède un gène létal qui n'est plus exprimé lors de la ligation de l'insert dans le site de clonage. Seules les cellules qui ont intégré le vecteur avec insert sont capables de croître sur le milieu de culture.

5.6.5 Clonage sans ligase (Polymerase Incomplete Primer Extension, PIPE)

Cette méthode de clonage a été décrite par Klock *et al.* en 2008 (Klock *et al.*, 2008). Son principe s'appuie d'une part sur la faculté qu'ont des extrémités d'ADN simple brin complémentaires à s'auto appairer, et d'autre part sur la capacité des cellules bactériennes à réparer l'ADN simple brin. Klock et ses collaborateurs ont montré que lors de l'amplification d'un fragment d'ADN, l'élongation n'est pas complète, et donc des extrémités 3' simple brins se forment de part et d'autre du produit de PCR. Dans son principe, le clonage PIPE consiste à amplifier séparément l'ADN à cloner et le plasmide avec des amorces comportant des séquences complémentaires (Figure III.3). Ainsi, le produit PCR de l'insert présentera des extrémités 5' compatibles avec celles du vecteur permettant l'auto-assemblage des deux molécules. Afin que l'assemblage soit stable, la zone d'appariement des séquences doit être d'au moins 15 bases. Lors de la transformation, les cellules hôte réparent les zones d'ADN simple brin et complètent la ligation vecteur-insert.

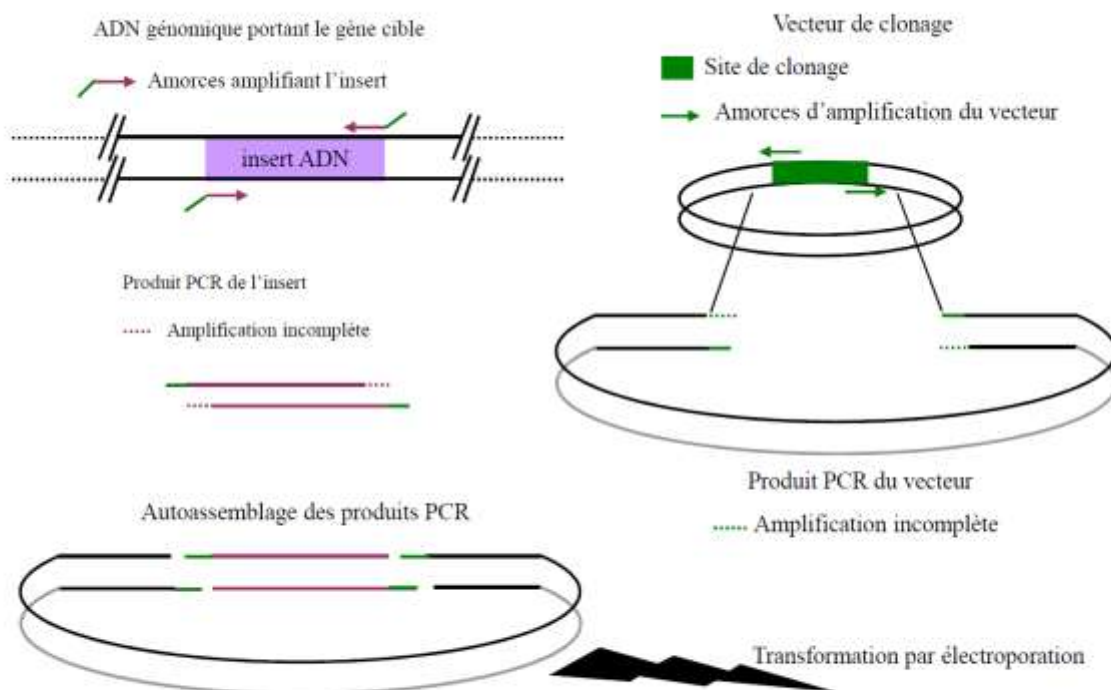


Figure III.3 : Principe du clonage PIPE. Le fragment d'ADN à insérer ainsi que le vecteur de clonage sont amplifiés par PCR. Les amorces utilisées pour amplifier l'insert présentent à leurs extrémités 5' quinze bases complémentaires aux extrémités du vecteur (vert) et quinze bases complémentaires à l'insert (violet).

Le clonage PIPE a été utilisé pour insérer les domaines catalytiques de dioxygénase amplifiés à partir de l'ADN extrait du sol dans le vecteur pFMA1.

Les domaines catalytiques ont été amplifiés avec les amorces répertoriés en annexe 1 (dcatdiox). Une partie de leur séquence (15 pb) est complémentaire de celles des amorces du vecteur. Le vecteur pFMA1 (4982 pb) a été amplifié avec les amorces PhnRHDr-V/PhnRHdf-V. Les conditions PCR sont résumées dans le tableau des amorces en annexe 1.

Afin d'améliorer l'efficacité du clonage, nous avons abaissé la quantité de vecteur matrice dans la réaction de PCR (< 5 ng), et digéré le produit PCR avec DpnI, enzyme qui coupe exclusivement l'ADN méthylé (le vecteur matrice). Ceci a permis de réduire le bruit de fond dû au vecteur utilisé comme matrice.

Les produits PCR ont ensuite été extraits par un volume de chloroforme : alcool isoamylique (24 : 1) et précipités à l'éthanol, en présence d'acétate de sodium (0,3 M) à - 20°C. Chaque amplicon a été repris séparément dans un volume d'eau stérile, dosé au NanoDrop avant d'être mélangés dans un ratio molaire de 1:3 (vecteur : insert). Le mélange a été transformé dans des cellules Top10 électro-compétentes. De multiples essais ont montré que les Top10 électro-compétentes, donnaient les meilleurs résultats en termes d'efficacité de clonage.

5.7 Vérification des clonages

Lors de clonage dans les vecteurs portant le gène lacZ (pUC18, pSTBlue), un criblage blanc/bleu en présence de X-gal (stock 20 mg/ml dans le DMSO ; 1/500^e dans le LB agar) a été réalisé. Les clones ayant incorporé l'insert forment des colonies blanches

Dans tous les cas, les clones sélectionnés ont été criblés par PCR. Lorsque l'insert a été détecté, les colonies ont été repiquées dans un milieu LB liquide avec l'antibiotique adéquate.

Les plasmides ont ensuite été purifiés, comme décrit précédemment (paragraphe 5.5 de Matériel et Méthodes) puis digérés par des enzymes de restriction. Les sites de restriction ont été choisis de manière à vérifier l'insertion de l'insert, son orientation et la taille du fragment inséré (ces sites de restriction seront détaillés dans les chapitres de résultat)

Enfin, les constructions ont été vérifiées par séquençage réalisé par Cogenics (Beckman) ou par Eurofins.

5.8 Profils de diversité

5.8.1 Polymorphisme de longueur des fragments de restriction (Restriction Fragment Length Polymorphism, RFLP)

Les gènes à analyser ont été amplifiés selon de protocole de l'enzyme GoTaq[®] DNA polymérase dans 12,5 µl de volume réactionnel. Les produits d'amplification ont été vérifiés sur gel d'agarose (4,5 µl). Les amplicons présentant la taille attendue ont été digérés par l'enzyme de restriction AluI (Fermentas), 1 h à 37°C. Cette enzyme possède une haute fréquence de coupure dans l'ADN. Les produits de digestion ont été analysés sur un gel de polyacrylamide (10%) préparé selon le protocole suivant : 5,46 ml H₂O, 3,33 ml acrylamide/bis acrylamide 30% (29 : 1) (Sigma), 1 ml tampon TBE (10x) (solution stock TrisHCl (90 mM), Acide Borique (90 mM), EDTA (2 mM)), 200 µl APS (10% w/v), 10 µl TEMED (Sigma). L'électrophorèse a été réalisée dans le tampon TBE (1x) à 200V, 1 h. Le profil de restriction a été révélé par coloration au GelRedTM.

5.8.2 Polymorphisme de conformation des simples brins d'ADN (Single Strand Conformation Polymorphism, SSCP)

Les analyses par SSCP ont été réalisées sur des échantillons de microcosmes incubés en présence ou non de phénanthrène. Lors de chaque analyse, 1 g de sol a été prélevé et divisé en échantillons de 250 mg. Les analyses ont été réalisées sur 6 à 9 réplicats selon les cas. L'ADN génomique a été extrait de 250 mg de sol suivant la méthode du kit PowerSoilTM DNA Isolation Kit. Les échantillons d'ADN génomique, ajustés à 10 ng/µl, ont servi de matrice à l'amplification de la région variable V3 (Figure I.4) du gène d'ARN ribosomal 16S (ARNr 16S) à l'aide des amorces W49fam et W104r (tableau des amorces – Annexe 1). L'amorce W49fam a été marquée en 5' par le phosphoramidite de fluoresceine. Le mélange réactionnel contenait 10 ng d'ADN génomique, 2,5 µl de tampon 10x, 2,5 µl MgCl₂ (25 mM), 0,5 µl dNTP (2,5 mM de chaque dNTP), 1 µl de chaque amorce (5 µM chacune), 0,2 µl BSA (20 g/l), 1 unité de AmpliTaq GoldTM DNA polymérase (Applied Biosystems Courtaboeuf, France) dans un volume total de 25 µl. Après amplification, 3 µl de chaque produit PCR a été analysé sur gel d'agarose 1,5% afin de vérifier la présence de la bande attendue à 205 pb. Les analyses SSCP par électrophorèse capillaire (CE-SSCP) ont été réalisées sur un ABI PRISM

3100 Genetic Analyser (Applied Biosystems) au Laboratoire d'Ecologie Alpine (LECA, Université de Grenoble) selon une méthode utilisée en routine (Zinger *et al.*, 2007).

Les données brutes issues du logiciel Genetic Analyser (Applied Biosystems) ont été représentées sous forme de graphiques de fluorescence non normalisés. Le logiciel (LucieAnalyzer) interne au LECA a été utilisé pour convertir ces graphiques en tableaux donnant les valeurs de fluorescence en fonction du temps de migration, converti en poids moléculaire. Les données de fluorescence ont ensuite été normalisées à l'aide du logiciel SCILAB (Scripts SSCP – Annexe 2), puis soumises à une analyse statistique à l'aide de l'application Vegan du logiciel R. Les distances entre les profils ont été obtenues en appliquant la fonction NMDS (nonmetric multidimensional scaling). L'influence du temps d'incubation (comme variable continue) et de la teneur en phénanthrène (comme variable ponctuelle) sur les profils SSCP a été évaluée grâce à la fonction *envifit* de Vegan (Scripts SSCP – Annexe 2).

5.9 Construction et séquençage des librairies de gènes d'ARNr 16S

Des échantillons d'ADN issus de sol ou des expériences SIP ont été utilisés comme matrice pour amplifier les gènes d'ARNr16S bactériens. Ces amplifications ont été réalisées à l'Institut de Génomique (Génoscope, Evry) avec les amorces 8F et 1390R, permettant d'obtenir la séquence presque complète des gènes (1382 pb). Les produits PCR ont été clonés dans le vecteur pCR™4 TOPO® TA (Invitrogen) pour constituer une librairie de clones par échantillon d'ADN. Le séquençage bidirectionnel Sanger a été réalisé selon un protocole standard. Les séquences ont été assemblées par PHRAP (<http://www.phrap.org>). Elles ont été soumises aux banques de données DDBJ/EMBL/GenBank sous les numéros d'accès : FQ658499 à FQ660546 et FQ790244

Les séquences de gènes d'ARNr 16S délivrées par le Génoscope ont été traitées par deux méthodes complémentaires. Dans un premier temps, les séquences de chacune des 9 librairies ont été affiliées à leur plus proche taxon bactérien à l'aide du programme SeqMatch disponible sur le site Ribosomal Database Project (<http://rdp.cme.msu.edu/index.jsp> ; (Cole *et al.*, 2009)). Les séquences ont été regroupées par ensembles taxonomiques (classe, ordre,

famille ou genre), et le nombre obtenu pour chaque ensemble a été rapporté au nombre total de séquences de la librairie.

Dans un second temps, l'ensemble des séquences des 9 librairies (2049 séquences) a été analysé *via* la suite d'applications mothur (Schloss *et al.*, 2009). Les séquences ont été alignées et regroupées en OTU (operational taxonomic units) en se basant sur un critère de dissimilarité de 3%. Les courbes de raréfaction relatives à chaque librairie ont été calculées, puis extrapolées jusqu'à l'asymptote au moyen d'une équation du type Michaelis-Menten. Les indices de diversité et les estimateurs de richesse ont été calculés pour chaque librairie indépendamment. Enfin, le nombre d'OTU partagés entre plusieurs librairies (jusqu'à 4) a été déterminé grâce à l'application Venn de mothur (Scripts mothur – Annexe 3).

5.10 Comparaison des domaines catalytiques de dioxygénases

L'obtention des séquences des domaines catalytiques de dioxygénases à partir d'ADN génomique marqué ^{13}C sera détaillée dans le chapitre résultats dédié à ce projet. Les séquences nucléotidiques ont tout d'abord été traduites *via* l'exécutable ApE (A plasmid Editor : <http://biologylabs.utah.edu/jorgensen/wayned/appe/>). Dans un premier temps, les séquences d'acides aminés ne présentant pas de codon STOP ont été comparées aux bases de données NCBI (National Center for Biotechnology Information) *via* blastp (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Dans un second temps, les séquences expérimentales ont été alignées, *via* clustalX, avec des séquences issues des bases de données. Toutes ces séquences ont été tronquées afin d'obtenir uniquement le domaine catalytique. Un arbre phylogénétique a été créé *via* le logiciel en ligne Phylogeny.fr (Gascuel *et al.*, 2008).

6. PCR quantitative en temps réel (qPCR)

6.1 Principe de la méthode

La qPCR a pour objectif de quantifier, dans un échantillon d'ADN, un gène spécifique. Par extension, il est possible de quantifier des groupes phylogénétiques spécifiques *via* leur gène d'ARNr 16S. L'amplification d'un gène par PCR suit une courbe sigmoïde comme présenté Figure III.4-A. Si la quantité initiale de matrice est faible, la phase plateau sera atteinte à l'issue d'un plus grand nombre de cycles, mais la valeur du plateau sera équivalente. Ainsi,

lors d'une PCR classique à 30 ou 35 cycles, le produit final analysé sur gel ne peut renseigner sur la quantité de matrice initiale puisque le plateau d'amplification est atteint. En revanche, lors de la phase exponentielle, la variation du nombre d'amplicons à chaque cycle dépend directement de la quantité initiale de matrice (Figure III.4-B). Quand l'efficacité d'amplification est de 100%, on peut calculer le nombre de copies d'un gène obtenu après « c » cycles par la formule : $N_C = N_0 \times 2^C$ où « N_0 » représente la quantité initiale de matrice.

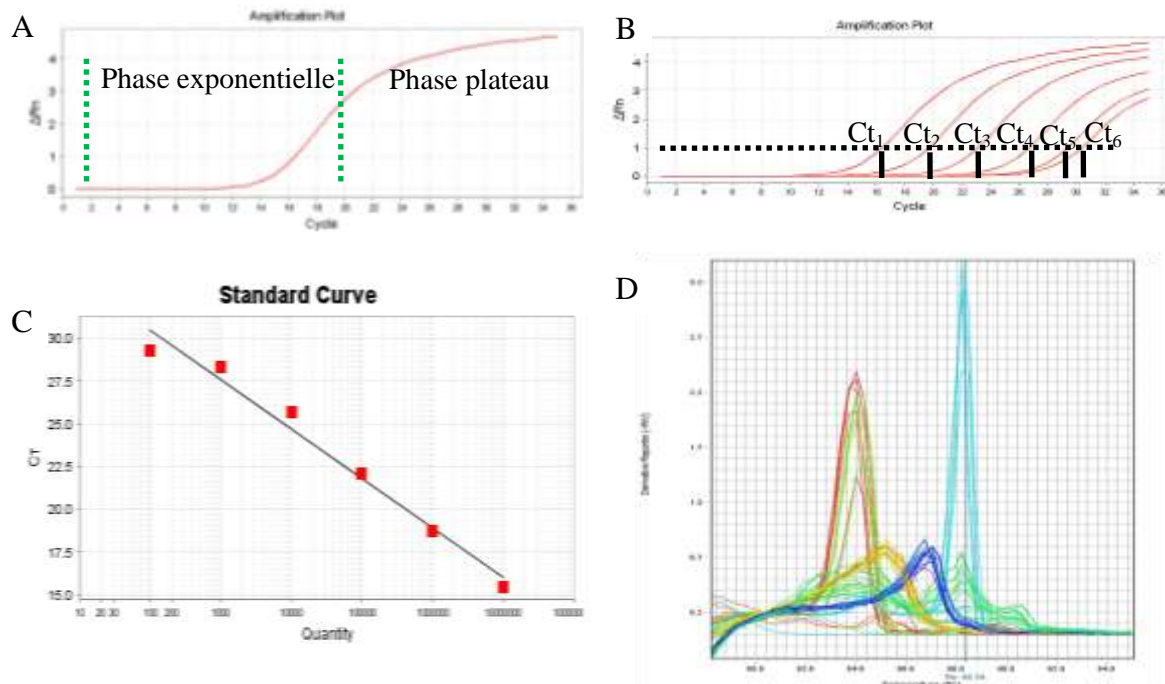


Figure III.4 : A : courbe d'amplification par PCR. La sigmoïde se divise en deux phases : la phase exponentielle peut être modélisée par une fonction qui renvoie à la quantité de produit synthétisé à chaque cycle en fonction de la quantité initiale de matrice. La phase plateau correspond à une chute de l'efficacité d'amplification (épuisement des réactifs) ; elle n'est pas modélisable et n'est pas toujours reproductible. B : courbes d'amplification d'un gène selon une gamme de dilutions de la matrice de départ. La définition d'une valeur seuil de fluorescence permet de calculer le nombre de cycles C_t caractéristique de l'échantillon. C : gamme étalon obtenue en reportant les C_t en fonction du nombre de copies du gène servant de matrice. D : courbes de fusion des amplicons. Chaque couleur correspond à un échantillon différent. Un pic unique prouve la spécificité de l'amplification.

Il existe plusieurs techniques pour détecter l'ADN amplifié en temps réel. Seule la technique SyBr Green, utilisée pour ces travaux, sera décrite ici. Le SyBr Green est une molécule qui, en s'intercalant dans les doubles brins d'ADN, forme un complexe beaucoup plus fluorescent que lorsqu'il est seul. A chaque cycle de PCR, l'appareil enregistre l'intensité du signal fluorescent de chaque échantillon. Une valeur seuil de fluorescence comprise entre le bruit de fond et la valeur plateau est déterminée à partir de la courbe sigmoïde obtenue. Le nombre de

cycles correspondant à cette valeur seuil de fluorescence est appelée le Ct. La quantification d'un gène dans un échantillon inconnu est réalisée de manière indirecte à partir d'une gamme étalon reliant la concentration initiale en gène cible au Ct correspondant (Figure III.4-C). A l'issue de l'amplification, un gradient de température est appliqué à tous les échantillons afin de déterminer la température de fusion des fragments synthétisés. Le nombre de pics obtenus sur la courbe de dénaturation (Figure III.4-D) correspond au nombre d'amplicons différents présents dans l'échantillon. Les paramètres d'amplification doivent être optimisés pour qu'un seul fragment soit amplifié dans chaque échantillon.

6.2 Tests d'inhibition

Les préparations d'ADN génomique issu du sol peuvent contenir des substances qui inhibent les ADN polymérases, comme les acides humiques par exemple. Afin de s'assurer qu'aucun biais n'est apporté par la qualité de l'extrait d'ADN, un test d'inhibition a été réalisé. Un vecteur de clonage (pGEM) apporté en quantité connue seul ou dans un échantillon d'ADN génomique du sol (2 ng/μl) a été quantifié par qPCR à l'aide de ses amorces spécifiques. Si les courbes d'amplification sont équivalentes en présence ou en absence d'échantillon environnemental, on en conclut que l'ADN ne contient pas d'inhibiteur à une concentration telle qu'il puisse biaiser la réaction de qPCR. Les tests d'inhibition des expériences de qPCR seront présentés dans l'annexe 4.

6.3 Quantification

Pour chaque couple d'amorces (Tableau des amorces – Annexe 1), une gamme étalon a été construite à partir de concentrations connues (10^7 à 10^3 copies par ng) du gène cible porté par le plasmide pGEM-T linéarisé. Les échantillons ont été dilués à 2 ng/μl. Le mélange PCR contenait 4 ng d'ADN génomique de sol, 7,5 μl de Master Mix SYBRGreen[®] PCR (Absolute QPCR SYBRGreen Rox, Abgene), 250 ng de gène 32 du phage T4 (QBiogene), 1,5 μl de chaque amorce 10 μM (Tableau des amorces – Annexe 1) dans un volume final de 15 μl. Les programmes d'amplification sont reportés dans le Tableau des amorces (Annexe 1). L'étape à 80°C permet de dénaturer tous les doubles brins non spécifiques (ne correspondant pas au fragment ciblé) afin de ne détecter que la fluorescence du gène cible. Cette température a été déterminée lors de la mise au point du protocole à partir des courbes de dénaturation.

L'amplification a été réalisée avec un appareil StepOnePlus™ Real-Time PCR Systems (Applied Biosystems).

Avec les données obtenues, une régression linéaire a été calculée entre les Ct et les concentrations de la gamme étalon. Le Ct mesuré pour les échantillons a été converti en nombre de copies.

6.4 Traitement des résultats de qPCR

Les résultats bruts générés lors d'une expérience de qPCR donnent le nombre de copies du gène cible contenu dans l'ADN analysé. Dans un premier temps, cette valeur a été rapportée à la quantité (en nanogrammes) d'ADN génomique qui a servi de matrice. La moyenne des réplicats donne une valeur de l'abondance du gène ciblé dans chaque échantillon (exprimé en nombre de copies/ng d'ADN). L'abondance relative de chaque phylum a été calculée en rapportant le nombre de copies de gènes ARNr16S du phylum considéré au nombre de copies de gènes ARNr 16S total (exprimé en %). Le test de Student a été appliqué pour savoir si les variations observées étaient significatives.

7. Méthodes de biochimie

7.1 Expression de dioxygénases recombinantes

Les gènes de dioxygénases clonés dans le vecteur pUC18 ont été co-exprimés dans *E. coli* avec les gènes codant des transporteurs d'électrons portés par un autre plasmide, pIBA34 (Tableau III.4)

Une culture liquide de 5 ml de LB a été inoculée par une colonie, puis incubée à 37°C jusqu'à atteindre une $DO_{600} = 1$. L'expression des protéines recombinantes a été induite par 0,5 ou 1 mM d'IPTG pendant 16 h à 25°C. L'induction a été contrôlée par analyse SDS-PAGE d'échantillons de suspension bactérienne prélevés avant et après induction. Les échantillons bactériens (culots après centrifugation) ont été lysés dans un tampon TEp (Tris HCl 50 mM pH 8 ou 7,5 selon les cas – EDTA 10 mM) 10 min à 95°C en présence de SDS Mix (1x) puis centrifugés 10 min à 16 200 $\times g$ avant électrophorèse sur gel de polyacrylamide à 12%.

La solution stock de SDS Mix (4x) est composée de Tris HCl (0,125 M ; pH 6,8), de beta mercapto éthanol (10%), de SDS (40 g/l), de glycérol (40%), et de Bleu de Bromophénol (0,004%) dans H₂O

7.2 Analyse SDS PAGE

Les protéines surexprimées dans *E.coli* ou extraite d'échantillons de sol (paragraphe 8.1 de Matériel et Méthodes) peuvent être visualisées sur gel de polyacrylamide en présence du détergent SDS qui permet d'uniformiser la charge des protéines. Celles-ci, ne portant que des charges négatives, vont migrer vers l'électrode positive du système d'électrophorèse (SE 260 Mighty Small II, basic, Hoefer™). Les électrophorèses ont été réalisées à l'aide de gel de 8x9 cm et 1 mm d'épaisseur dans un tampon de migration (Tris 25 mM, glycine 0,19 M, SDS 3,5 mM).

Les échantillons protéiques ont tout d'abord été déposés dans les puits d'un gel de concentration (0,75 ml H₂O – 0,3 ml solution III – 0,2 ml solution I – 25 µl APS 10% – 3 µl TEMED) puis soumis à un champ électrique de 10 à 20 mA. Par la suite, un gel de séparation, contenant 12% d'acrylamide (2,4 ml H₂O – 1,8 ml solution II – 3 ml solution I – 50 µl APS 10% – 10 µl TEMED) a permis de séparer les protéines, soumises à un champ électrique de 50 mA, selon leur masse moléculaire (Laemmli, 1970).

Après électrophorèse, le gel a été rincé à l'eau puis coloré au bleu de Coomassie (2 g Coomassie® Brilliant Blue R 250, 450 ml éthanol, 90 ml acide acétique, 450 ml H₂O) 10 min et décoloré dans une solution d'Ethanol (5%), d'acide acétique (7%) et H₂O

Solution I : Acrylamide/Bisacrylamide 37,5/1 (Euromedex)

Solution II : Tris HCl 1,5 M (pH 8,8), SDS 10 mM

Solution III : Tris HCl 0,25 M (pH 6,8), SDS 6,9 mM

7.3 Dosages *in vivo* de l'activité catalytique des dioxygénases

Les cultures de souches transformées avec les gènes de dioxygénases et induites par IPTG ont été centrifugées (2 500xg 10 min). Les culots ont été lavés par 5 ml de milieu M9 puis repris dans un volume de milieu M9 complet (4.2.1.2) adéquat pour ajuster la DO₆₀₀ à 2. Des réplicats de 1 ml de suspension bactérienne dans le milieu M9 ont été transférés dans des tubes eppendorf de 2 ml ayant reçu un des HAP (Tableau III.6) en solution dans l'acétone. Le solvant a été évaporé avant le transfert de la suspension. Les cultures ont été incubées à 25°C et 200 rpm pendant 6 h.

Tableau III.6 : HAP utilisés lors des dosages de l'activité des dioxygénases *in vivo*.

HAP	Concentration dans l'acétone	Concentration finale dans la culture	Temps d'incubation
Naphtalène	25 mM	1,25 mM	6 h
Biphényle	10 mM	0,5 mM	6 h
Phénanthrène	10 mM	0,5 mM	6 h
Anthracène	10 mM	0,5 mM	6 h

Les cultures ont été centrifugées 10 min à 6 200xg. Le surnageant (900 µl) a été transféré dans un nouveau tube eppendorf de 2 ml. Un étalon interne, le 2,3-dihydroxybiphényle a été ajouté au surnageant à une concentration finale de 10 µM, 5 µM ou 0,2 µM selon les cas. Les dihydrodiols ont été extraits par un volume équivalent d'acétate d'éthyle. La phase organique a été séchée sur sulfate de sodium puis le solvant a été évaporé. Les résidus secs ont été dissous dans 200 µl d'acétonitrile, et les diols ont été dérivatisés par du n-butyl boronate (NBB); 100 µl de solution diol a été incubé avec 50 µl de réactif NNB* (5 mg/ml) pendant 6 h.

Des gammes de calibration ont été préparées à partir de solutions de dihydrodiols de trois HAP, naphtalène, phénanthrène et biphényle, purifiés au laboratoire (Jouanneau and Meyer, 2006), diluées à des concentrations connues et dérivatisés de la même manière que les échantillons. Après centrifugation, les produits dérivatisés ont été analysés par GC/MS avec l'appareillage décrit précédemment (paragraphe 3.1.2 de Matériel et Méthodes).

Les échantillons (2,5 µl) ont été injectés sur une colonne Supelco MDN12. Le détecteur de masse HP 5973 réglé en mode SIM a détecté les complexes diols-NBB élués entre 8 et 25,8 min selon leur masse moléculaire (M^+), en ciblant les m/z suivants : naphthalène-1,2-dihydrodiol : 228 ; biphenyle-1,2-dihydrodiol : 254 ; phénanthrène-3,4-dihydrodiol : 278 ; anthracène-1,2-dihydrodiol : 278. Les deux derniers diols se distinguent par leur temps de rétention sur la colonne.

Les diols produits ont été quantifiés indirectement par le biais de la gamme de calibration. L'étalon interne a permis de corriger les mesures pour tenir compte des écarts dus aux manipulations (rendements d'extraction et de dérivatisation)

* Le réactif NBB (acide N-butyl boronique) est préparé en ajoutant 50 mg d'acide N-butyl boronique à un mélange de chloroforme (5 ml) et de 2,2-diméthoxypropane.

7.4 Purification des protéines recombinantes

7.4.1 Expression des dioxygénases hybrides

L'expression des dioxygénases recombinantes a été effectuée dans la souche *E.coli* JM109. Cette souche a été co-transformée par un vecteur dérivé de pUC18 portant les gènes de la dioxygénase à purifier, et par le plasmide pIBA34 portant les gènes des transporteurs d'électrons (pIBA34 n'est pas nécessaire pour la purification des dioxygénases). Les transformants ont été incubés dans 25 ml de milieu LB 16 h à 37°C et 200 rpm, puis ces précultures ont servi à inoculer des cultures de 800 ml de milieu LB. Lorsque la densité bactérienne a atteint une DO_{600} proche de 1, les cultures ont été induites avec 1 mM d'IPTG (solution stock 0,2 M) et supplémentées avec 50 µM de $Fe(NH_4)_2(SO_4)_2$. L'incubation s'est poursuivie à 20°C, 200 rpm pendant 22h, puis les cultures ont été centrifugées à 2 392xg 15 min (Avanti™ Centrifuge J-30I Beckman Coulter équipée du rotor JLA 9.1000). Après lavage dans 40 ml de phosphate de potassium 50 mM (pH 7,5), les culots ont été stockés à -20°C dans des tubes Falcon 50 ml.

7.4.2 Purification des dioxygénases recombinantes

Les tampons utilisés pour la purification (listés ci-dessous) ont été placés dans une boîte à gants sous argon pendant 20 h afin d'éliminer l'oxygène pour prévenir l'oxydation de l'ion ferreux du site catalytique des dioxygénases. Juste avant utilisation, 1 mM final de beta-mercaptoéthanol (solution stock 14 M) a été ajouté dans tous les tampons.

Tampons A : K_2HPO_4 25 mM, pH 7,5 – éthanol 5% – glycérol 5%

Tampon B : tampon A – NaCl 0,5 M

Tampon C : tampon A – NaCl 0,5 M – Imidazole 10 mM

Tampon E : tampon A – Imidazole 150 mM

Les culots bactériens ont été décongelés dans la boîte à gants et repris dans 20 ml de tampon B. Une fois homogénéisée, la suspension a été traitée au lysozyme (0,5 mg/ml final) pendant 15 min. Les cellules bactériennes ont été lysées par ultrasons à 0-4°C à l'aide d'un appareil Vibracell™ 72412 (Fisher Bioblock Scientific, Illkirch, France) équipé d'une microsonde. Des cycles de traitement par les ultrasons ont été réalisés à 30% de la puissance maximale pendant 5 min, avec une alternance de 5 s de pulse et 5 s de pause. Les extraits bruts ont été centrifugés 30 min à 15 500 $\times g$ à 4°C. Les lysats clarifiés ont été transférés dans de nouveaux tubes Falcon de 50 ml et maintenus à 0-4°C sous atmosphère d'argon.

Les protéines surproduites possèdent une extrémité poly-Histidine (His-tag) du coté N-terminal. Elles peuvent être purifiées par chromatographie d'affinité sur une résine sur laquelle des métaux divalents comme le cobalt, le nickel ou le zinc ont été fixés. Les protéines recombinantes ont été purifiées sur une résine d'affinité de type TALON (Clontech) sur laquelle est fixé du cobalt. Des seringues de 2 ml ont été remplies avec 1,5 ml de résine humide et les mini-colonnes ont été équilibrées avec 15 ml de tampon B. Les lysats ont été déposés sur les colonnes au moyen d'une pompe péristaltique multicanaux (Gilson). La colonne a ensuite été lavée avec 20 ml de tampon B, puis 10 ml de tampon C. Enfin, la protéine a été éluée dans le tampon E en fractions de 1 ml collectées dans des tubes à hémolyse.

7.4.3 Caractéristiques spectrales des protéines surproduites

Les fractions contenant la protéine purifiée ont été déterminées par mesure d'absorbance entre 250 nm et 700 nm (Diode Array Spectrophotometer Hewlett Packard 8452A). Les pics caractéristiques, des protéines (280 nm), du centre [2Fe-2S] (320 nm ; 420 nm et 460 nm) et des acides nucléiques (260 nm) renseignent sur la présence de la protéine, la conservation, et l'état d'oxydation des centres [2Fe-2S] ainsi que sur l'éventuelle contamination par l'ADN. Lorsque la protéine est oxydée, il est possible d'estimer sa concentration grâce à l'absorbance à 460 nm par la formule $C = (\epsilon_{460} \cdot l) / A_{460}$, avec C la concentration exprimée en mol/l (M), ϵ_{460} le coefficient d'extinction molaire à 460 nm exprimé en $\text{cm}^{-1} \cdot \text{M}^{-1}$ et l le chemin optique (1 cm). Les fractions contenant la protéine purifiée ont été concentrées dans des microcentrifugal concentrator (Microstep – seuil de coupure 30 kDa) préalablement équilibrés avec le tampon A. Les solutions protéiques de 500 μl environ ont été congelées sous forme de bille par l'azote liquide.

7.4.4 Dosage des protéines

Les protéines ont été dialysées 3 h contre le tampon 1 de dialyse (K_2HPO_4 (25 mM pH=7,5) – 5% Glycérol – 5% éthanol) puis 16 h contre le tampon 2 de dialyse (K_2HPO_4 (25 mM pH=7,5) – 5% Glycérol).

Le dosage des protéines a été réalisé selon la méthode BCA (Bicinchoninic Acid) suivant les instructions du Kit Pierce Protein Assay Reagent – Enhance protocole (Thermo Scientific).

Le dosage BCA (Smith *et al.*, 1985) repose sur la réaction de biuret. En conditions alcalines, les protéines chélatent le cuivre (Cu^{2+}) pour former l'ion cuivrique (Cu^+). L'intensité de la couleur bleue obtenue est proportionnelle au nombre de liaisons peptidiques impliquées et peut être mesurée à 540 nm.

Lorsque le l'acide bicinchoninique (BCA) se trouve dans la solution, le complexe BCA/Cu^+ qui se forme présente une forte absorbance à 562 nm qui augmente linéairement avec la concentration des protéines en solution. La détection du complexe violet BCA/Cu^+ est cent fois plus sensible que celle du complexe formé par la réaction du biuret.

La concentration des protéines d'une solution inconnue est calculée *via* une gamme étalon de BSA (sérum d'albumine bovine).

Le kit utilisé est composé de deux solutions :

Réactif A : carbonate de sodium, bicarbonate de sodium, acide bicinchoninic et tartrate de sodium dans 0,1 M de soude.

Réactif B : solution à 4% de sulfate de cuivre.

La solution de travail est préparée en mélangeant 50 volumes de réactif A avec 1 volume de réactif B.

Les échantillons de concentration inconnue et de la gamme étalon (50 µl) ont été incubés avec le réactif de travail (1 ml), 30 min à 60°C, puis refroidis. L'absorbance a été mesurée à 562 nm.

8. Extraction et détection des protéines issues des échantillons de sol

8.1 Extraction des bactéries du sol

Afin de contrôler le rendement des étapes d'extraction des protéines du sol, une souche de référence, *Pseudomonas* sp. KT2442 contenant le plasmide pSD9 portant les gènes de la dioxygénase de *Sphingomonas* sp. CHY-1 (Jouanneau *et al.*, 2006), a été introduite dans les échantillons de sol. L'expression des protéines recombinantes portées par les vecteurs a été induite par 1 mM d'IPTG. La surexpression a été effectuée à 25°C pendant 16 h. La suspension bactérienne a ensuite été ajustée à $DO_{600} = 1$ avant utilisation. Lors du suivi de l'efficacité des méthodes d'extraction, les échantillons issus des différents traitements, en présence de *Pseudomonas* sp. KT2442 ont étéensemencés sur des milieux solide LB supplémentés de tétracycline, dont la résistance est portée par le vecteur que contient la souche.

Six tampons différents ont été testés lors de la mise au point du protocole :

Tampon pyrophosphate de sodium 0,05 M pH 6,4

Tampon pyrophosphate de sodium 0,1 M pH 6,4 – pH 4 – pH 9

Solution saline $H_2O.NaCl$ 0,8%

Le détail de la mise au point du protocole sera présenté dans IV.Chapitre V.1.

Les détails techniques de chaque étape seront présentés ici.

De manière générale, 10 g humides de sol ont été mis en suspension dans 20 ml de tampon pyrophosphate de sodium ou de solution saline ($\text{H}_2\text{O.NaCl}$ 0,8%). Lorsque des lavages ont été réalisés dans les tampons pyrophosphate de sodium 0,1 M pH 6,4, 4 ou 9, les échantillons de sols ont été agités au vortex puis centrifugés à 10 min 10 000 xg .

Afin de décoller les bactéries des particules de sol, un traitement mécanique a été appliqué, soit par ultrasons soit par broyage. Le traitement par ultrasons a été réalisé avec le même appareil que décrit précédemment (paragraphe 7.4.2) à une puissance de 20% la puissance maximale pendant 30 s. Le broyage a été effectué dans un mixeur (Waring Blender[®]) à vitesse maximale pendant une minute (bol de 350 ml). Quand plusieurs cycles de broyage ont été nécessaires, ils ont été séparés de 5 min de pause dans la glace. Cette étape de traitement mécanique a parfois été associée à un traitement chimique, par ajout de surfactant tels que le Tween20 (0,5%), le Triton(x100) (0,5%) ou le PVPP (polyvinyl polypyrrolidone) (0,5%).

La suspension homogène obtenue après ces traitements a été transférée (20 ml) dans un tube d'ultracentrifugeuse Ultra-Clear (38,5 ml – Beckman Coulter). Un coussin de Nicodenz (AbCys S.A., Paris) (9,2 ml d'une solution stock 0,8 g/ml) a été ajouté en dessous de la suspension de sol. Une ultracentrifugation a été réalisée avec le rotor SW28 (Beckman Coulter) dans une ultracentrifugeuse L-70 (Beckman Coulter) 1 ou 2 h à 14 000 xg .

Après centrifugation, le halo bactérien à l'interface Nicodenz/phase aqueuse a été transféré, à l'aide d'une pipette de transfert, dans un tube Falcon de 50 ml. Plusieurs lavages de la suspension bactérienne ont été réalisés avec des volumes décroissants de tampon pyrophosphate (0,05 M pH 6,4) ou de solution saline. Le culot obtenu a été conservé à -20°C en attendant son traitement.

8.2 Extraction des protéines

Les bactéries isolées du sol ont été reprises dans un tampon Tris HCl (50 mM) pH 8 – EDTA (10 mM) contenant dans certains cas des agents de lyse (lysonase ou benzonase) ou dans un tampon de lyse (Urée 9 M – CHAPS 65 mM – Triton 0,5x – DTT 0,1 M). Les cellules ont été lysées soit par traitement aux ultrasons (10 % de l'amplitude maximale) pendant 1 min 30 s

avec des pulses de 9 s, soit par traitement à la soude (0,2 M) 1h à 90°C suivi d'une neutralisation dans le tampon MES (0,3 M final).

De même, différentes méthodes ont été testées pour isoler les protéines ; une précipitation acide, avec 10% d'acide trichloracétique (TCA – solution stock 75%) suivi d'une neutralisation au Tris Base, une extraction phénol dans un bain à ultrasons suivi soit d'une extraction par l'acétate d'ammonium 0,1 M dans le méthanol et d'un lavage à l'éthanol 70%, soit d'une extraction par le diisopropyléther à froid. Quelque soit la méthode utilisée, les échantillons ont été repris dans un volume adéquat de SDS Mix (1x) à 100°C puis analysés par SDS PAGE sur un gel 12%.

Lorsque des protéines ont été détectées, elles ont été analysées par chromatographie liquide et spectrométrie de masse comme décrit au paragraphe 3.4 de Matériel et Méthodes.

IV. Résultats et Discussions

Chapitre I. Mise en place du cadre de l'étude : expérimentations préliminaires

Les hydrocarbures aromatiques polycycliques (HAP) sont des composés hydrophobes et toxiques pour l'environnement. Leur stabilité rend ces molécules très difficiles à oxyder tant chimiquement que par voie microbienne. Pourtant il existe des microorganismes capables d'initier l'attaque de ces cycles, voire de les utiliser comme sources uniques de carbone et d'énergie (I.2.4 de l'Etude Bibliographique).

Lors de ce projet, nous avons étudié l'impact de ces polluants sur les populations bactériennes d'un site alimenté de manière chronique par des hydrocarbures.

Au-delà de la caractérisation de la diversité bactérienne du site dans son ensemble, nous avons cherché à savoir (1) si cette diversité était affectée par un pic de pollution de type HAP, (2) quel sous-ensemble de cette population était directement impliqué dans la dégradation des HAP.

Afin d'identifier plus aisément les bactéries capables de métaboliser les HAP *in situ*, nous avons opté pour une approche de marquage isotopique décrite sous le vocable anglais de 'Stable Isotope Probing' (SIP).

La mise en œuvre de la méthode SIP nécessite

- de choisir un substrat représentatif de la famille des HAP et de le marquer uniformément avec un isotope stable,
- d'apporter ce substrat à une dose permettant à la fois de marquer spécifiquement les populations actives et d'éviter d'éventuels effets toxiques sur l'ensemble des populations bactériennes du sol. Cela permet de conserver le bénéfice des interactions au sein des communautés bactériennes,
- d'ajuster les conditions expérimentales afin d'éviter le marquage des populations non impliquées dans la dégradation du substrat (cross-feeding). Ces populations sont constituées d'organismes saprophytes ou de bactéries utilisant des métabolites du substrat marqué.

Ce chapitre présente la synthèse du substrat uniformément marqué au ^{13}C , le dosage des HAP dans les sols d'étude ainsi que des incubations en microcosmes réalisées afin de déterminer les conditions expérimentales propices au marquage isotopique (SIP).

1. Synthèse du phénanthrène marqué au carbone ^{13}C

Nous avons choisi de mettre en œuvre l'approche SIP en utilisant le phénanthrène comme substrat modèle. C'est un HAP de trois cycles métabolisé par un grand nombre de bactéries (paragraphe I.2.6 de l'Etude Bibliographique).

Le phénanthrène uniformément marqué avec l'isotope ^{13}C du carbone a été synthétisé au laboratoire. Le mode opératoire de la synthèse (Singleton *et al.*, 2005) a été optimisé afin d'augmenter les rendements de chaque étape.

Des analyses par Résonnance Magnétique Nucléaire (RMN) et par Chromatographie Gazeuse et Spectrométrie de Masse (GC/MS) ont permis d'assurer que le produit obtenu était le phénanthrène et qu'il était effectivement marqué au ^{13}C .

1.1 Modes opératoires

La stratégie initiale a consisté à synthétiser le phénanthrène marqué *via* la synthèse du naphthalène à partir du benzène et de l'acide succinique uniformément marqués au ^{13}C , comme le décrit l'équipe de Singleton (Singleton *et al.*, 2005). La mise au point de la synthèse a été réalisée sur des molécules non marquées (^{12}C) (Figure IV.1).

1.1.1 Synthèse du naphthalène à partir du benzène

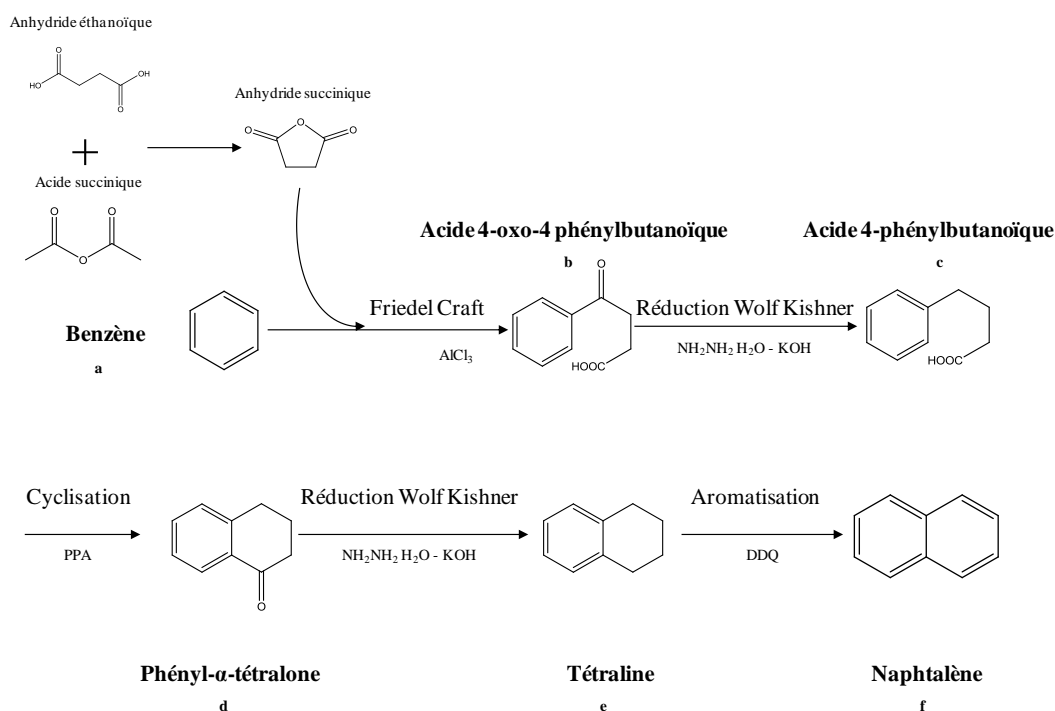


Figure IV.1 : Synthèse du naphthalène à partir du benzène et de l'acide succinique.

Synthèse de l'anhydride succinique

Un mélange d'acide succinique (5,01 mg, 118 g/mol, $4,24 \cdot 10^{-2}$ mol) et d'anhydride acétique (15 mL) est chauffé 2 h au reflux (150°C), puis refroidi. L'addition d'éther isopropylique provoque la formation d'un précipité blanc. Après filtration sur fritté et séchage au dessiccateur, le produit est isolé sous forme d'une poudre blanche ($m = 2.81$ g, $R = 66\%$).

Synthèse de l'acide 4-oxo-4 phénylbutanoïque (b)

Avant de commencer la réaction, le ballon est placé sous atmosphère inerte (argon) ; cette atmosphère est conservée durant la réaction. Un mélange de benzène (**a**) (2 g, 78g/mol, $2,56 \cdot 10^{-2}$ mol) et d'anhydride succinique (2,46 g, 100 g/mol, $2,46 \cdot 10^{-2}$ mol) dans 20 ml de 1,2-dichloroéthane est maintenu à 0°C (dans un cristalliseur contenant de l'eau et de la glace). AlCl_3 (8,44 g, 133 g/mol, $6,3 \cdot 10^{-2}$ mol) est ajouté à la suspension. Une fois que la température s'est stabilisée en dessous de 30°C , le mélange réactionnel est agité à température ambiante pendant 4 heures. La réaction est arrêtée en ajoutant 50 ml d'un mélange glace/eau. Le produit est extrait par du dichlorométhane (2 x 100 ml). L'ensemble des phases organiques est réuni et extrait avec HCl 1 M (1 x 100 ml) puis séché sur NaSO_4 . Le solvant est évaporé pour donner une poudre blanche de (**b**) ($m = 4$ g, $R = 93\%$).

Synthèse de l'acide 4-phénylbutanoïque (c)

Pour cette étape, seule une partie de **(b)** a été utilisée. Plusieurs essais ont été réalisés pour mettre au point la synthèse. Seule une réaction type sera présentée.

Un mélange de **(b)** (1,36 g, 166 g/mol, $3,53 \cdot 10^{-3}$ mol), de potasse (KOH, 2,859 g, 56 g/mol, $51 \cdot 10^{-3}$ mol) et d'hydrazine (NH_2NH_2 , 2,5 g, 50 g/mol, $50 \cdot 10^{-3}$ mol) dans 10 ml d'éthylène glycol est chauffé 6 h au reflux (150°C), puis refroidi à 80°C. La réaction est arrêtée par addition de 30 ml d'un mélange glace/eau, acidifiée à pH 1 avec HCl 12N et extraite au chloroforme. La phase organique est séchée sur MgSO_4 . Le solvant est évaporé pour donner un semi-solide orange-brun de **(c)** (m = 700 mg, R = 51%).

Synthèse de la phényl- α -tétralone (d)

700 mg de **(c)** sont mélangés à de l'acide polyphosphorique (PPA, 5,77 g, 335 g/mol, $17 \cdot 10^{-3}$ mol), chauffés au reflux 25 min (95°C). La réaction est refroidie avec un mélange glace/eau et le produit est extrait au diéthyléther (1 x 50 ml). La phase organique est lavée avec une solution saturée de NaHCO_3 puis séchées sur MgSO_4 . Le solvant est évaporé pour donner un résidu huileux orange-brun contenant **(d)** (m = 520 mg, R = 83%).

La réduction et la cyclisation ont été réalisées deux fois de suite sur la même proportion de produit afin d'améliorer les rendements. L'optimisation a porté sur le traitement du produit.

Les produits des deux cyclisations ont été réunis pour donner 1,25 g de **(d)**.

Synthèse de la tétraline (e)

Un mélange de **(d)** (660 mg, 146 g/mol, $4,5 \cdot 10^{-3}$ mol), de potasse (KOH, 1,52 g, 56 g/mol, $27 \cdot 10^{-3}$ mol) et d'hydrazine (NH_2NH_2 , ,36 g, 50g/mol, $27 \cdot 10^{-3}$ mol) dans 10 ml d'éthylène glycol est chauffé 6 h au reflux (150°C), puis refroidi à 80°C. La réaction est arrêtée par addition de 30 ml d'un mélange glace/eau, acidifié à pH 1 avec HCl 12N et extraite par le dichlorométhane (1 x 50 ml). La phase organique est séchée sur NaSO_4 . Le solvant est évaporé pour donner un résidu huileux orange-brun de **(e)** (m = 527 mg, R = 89%).

Synthèse du naphthalène (f)

Un mélange de **(e)** (194 mg, 130 g/mol, $1,47 \cdot 10^{-3}$ mol) et de dichlorodicyanobenzoquinone (DDQ, 350 mg, $1,5 \cdot 10^{-3}$ mol) dans 10 ml de benzène est chauffé au reflux 45 min (80°C). Un équivalent supplémentaire de DDQ est ajouté et la température est maintenue à 80°C pendant

75 min. Le mélange est refroidi, le solvant est évaporé, produisant une poudre qui est dissoute dans l'éther de pétrole. Le produit est déposé sur une colonne de chromatographie sur alumine puis élué dans l'éther de pétrole. Le solvant est évaporé pour obtenir une poudre blanche de **(f)** ($m = 13,2 \text{ mg}$, $R = 7\%$).

Les rendements de synthèse de chaque étape intermédiaire sont satisfaisants. Cependant, le rendement de la dernière étape est très faible. Le naphthalène étant un HAP très volatil, nous pensons qu'il a été perdu lors de l'évaporation sous vide du solvant. Pour éviter le risque d'une perte du naphthalène à cette étape, nous avons décidé de synthétiser le phénanthrène marqué à partir de ^{13}C -naphthalène et de ^{13}C -succinate commerciaux.

1.1.2 Synthèse du phénanthrène à partir du naphthalène

La synthèse du phénanthrène a été réalisée selon le schéma de la Figure IV.2. Les étapes ont été optimisées sur des composés non marqués avant de passer à la synthèse du phénanthrène à partir d'environ 500 mg de ^{13}C -naphthalène.

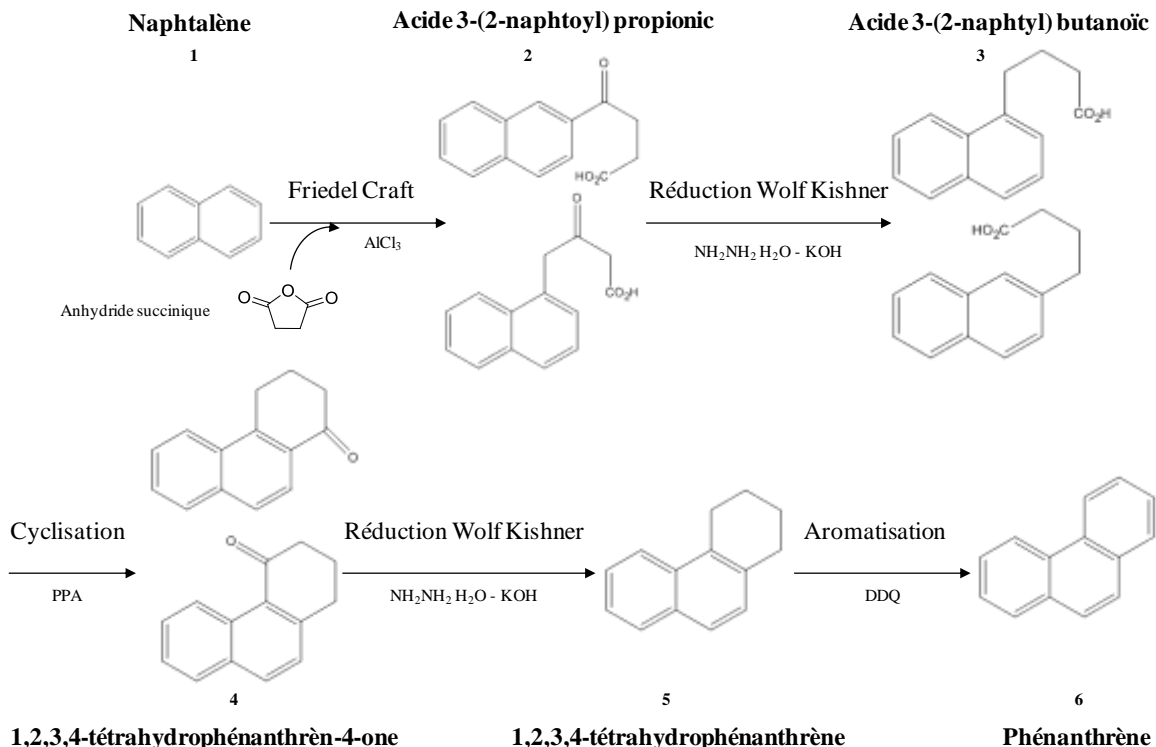


Figure IV.2 : Synthèse du phénanthrène à partir du naphthalène et de l'anhydride succinique.

Synthèse du [U-¹³C] acide 3-(2-naphtoyl) propionique (2)

Avant de commencer la réaction, le ballon est placé sous atmosphère inerte (argon) ; cette atmosphère est conservée durant la réaction. Un mélange de naphthalène (**1**) (511 mg, 138g/mol, $3,7 \cdot 10^{-3}$ mol) et d'anhydride succinique (385 mg, 104 g/mol, $3,7 \cdot 10^{-3}$ mol) dans 20 ml de 1,2-dichloroéthane est maintenu à 0°C (dans un cristallisoir contenant de l'eau et de la glace). AlCl₃ (1,23 g, 133 g/mol, $92,7 \cdot 10^{-3}$ mol) est ajouté à la suspension. Une fois que la température s'est stabilisée en dessous de 30°C, le mélange réactionnel est agité à température ambiante pendant 4 h. La solution est verte au départ et vire au vert noir. La réaction est arrêtée en ajoutant 50 ml d'un mélange glace/eau. Le produit est extrait au dichlorométhane (ou au chloroforme) (2 x 100 ml). Les phases organiques réunies sont extraites avec HCl 1 M (1 x 100 ml) et séchées sur MgSO₄. Le solvant est évaporé pour donner une poudre blanche de (**2**) (m = 855 mg, R = **95,5%**).

Synthèse du [U-¹³C] acide 3-(2-naphtyl) butanoïque (3)

Un mélange de (**2**) (855 mg, 242 g/mol, $3,53 \cdot 10^{-3}$ mol), de potasse (KOH, 1,187 g, 56 g/mol, $21,2 \cdot 10^{-3}$ mol) et d'hydrazine (NH₂NH₂, 1,06 g, 50 g/mol, $21,2 \cdot 10^{-3}$ mol) dans 10 ml d'éthylène glycol est chauffé 16 h au reflux (150°C), puis refroidi à 80°C. La réaction est arrêtée avec 30 ml d'un mélange glace/eau, acidifiée à pH 1 avec HCl 12N et extraite au dichlorométhane (2 x 100 ml). Les phases organiques combinées sont séchées sur MgSO₄. Le solvant est évaporé pour donner un résidu orange-brun visqueux de (**3**) (m = 838 mg, R = **98%**).

Synthèse du [U-¹³C] 1,2,3,4-tetrahydrophenanthren-4-one (4)

(**3**) est mélangé à de l'acide polyphosphorique (PPA, 9,85 g, 335 g/mol, $29,4 \cdot 10^{-3}$ mol), chauffé 25 min à 95°C. La réaction est refroidie avec un mélange glace/eau et le produit est extrait au diéther (4 x 50 ml). Les phases organiques combinées sont lavées avec une solution saturée de NaHCO₃ puis séchées sur NaSO₄. Le solvant est évaporé pour donner un résidu huileux orange-brun contenant (**4**) qui est utilisé sans purification.

Synthèse du [U-¹³C] tetrahydrophenanthrene (5)

Un mélange du produit (**4**), de potasse (KOH, 1,2 g, 56 g/mol, $21,5 \cdot 10^{-3}$ mol) et d'hydrazine (NH₂NH₂, 1,07 g, 50g/mol, $21,5 \cdot 10^{-3}$ mol) dans 10 ml d'éthylène glycol est chauffé 16 h au reflux (150°C), puis refroidi à 80°C. La réaction est arrêtée avec 30 ml d'un mélange

glace/eau, acidifié à pH 1 avec HCl 12N et extraite au dichlorométhane (3 x 100 ml). Les phases organiques combinées sont séchées sur NaSO₄. Le solvant est évaporé pour donner un résidu huileux orange-brun de **(5)** (m = 660 mg, R = **93,8%**)

Synthèse du [U-¹³C] phénanthrène (6)

Un mélange de **(5)** (659 mg, 196 g/mol, 3,37.10⁻³ mol) et de dichlorodicyanobenzoquinone (DDQ, 764 mg, 3,37.10⁻³ mol) dans 10 ml de benzène est chauffé 45 min au reflux (80°C). Un équivalent supplémentaire de DDQ est ajouté et la température est maintenue à 80°C pendant 75 min. Le mélange est refroidi, le solvant est évaporé, produisant une poudre noire qui est redissoute dans un mélange chloroforme/eau (200 ml, 1 : 1). Une chromatographie sur couche mince (CCM, silice) effectuée dans le chloroforme permet de s'assurer que la réaction est totale. La phase aqueuse est extraite au chloroforme (4 x 200 ml). Les phases organiques combinées sont extraites par l'eau (2 x 800 ml) et séchées sur NaSO₄. Le solvant est évaporé donnant une poudre marron. Ce solide est dissous dans un petit volume de dichlorométhane et déposé sur une colonne de chromatographie sur silice éluée par le chloroforme. Le produit co-élue avec une impureté. Le solvant est évaporé et la poudre jaune pâle produite est reprise dans 1 ml de chloroforme auquel sont ajoutés progressivement 5 ml de pentane. Une seconde colonne de chromatographie sur silice permet d'éluer, dans le mélange pentane/chloroforme (5 : 1), un produit pur (vérification par CCM, silice). Le solvant est évaporé pour obtenir une poudre blanche de **(6)** (m = 287,1 mg, R = **44%**).

Les rendements des cinq étapes sont nettement plus satisfaisants. Les rendements initialement obtenus pour ces étapes ont été optimisés en faisant varier les conditions de traitement des produits bruts de chaque réaction. Ainsi, le rendement total de la synthèse est passé de 19% à 40% (rendement obtenu pour la synthèse du produit marqué).

1.2 Analyse des produits de synthèse par Résonance Magnétique Nucléaire

1.2.1 Principes de RMN et matériel utilisé

La Résonance Magnétique Nucléaire (RMN) est une méthode d'analyse des composés chimiques qui donne la carte d'identité de la molécule étudiée. Les atomes qui peuvent être détectés par RMN possèdent un spin nucléaire non nul, autrement dit, un nombre impair de

protons et de neutrons. Ainsi, l'hydrogène (^1H), possédant 1 proton, et le ^{13}C , possédant 6 protons et 7 neutrons, sont les noyaux les plus couramment détectés par RMN avec les instruments classiques. Le ^{12}C (6 protons ; 6 neutrons) ayant un spin nucléaire nul, il ne peut être détecté par RMN.

En pratique, on distingue deux types d'appareillage : les spectromètres à onde continue ou à onde impulsionnelle. Dans le premier cas, on applique une fréquence fixe et le champ B est variable. Dans le second, le plus courant de nos jours, une brève radiofréquence est envoyée et le champ magnétique est fixe. La technologie à onde impulsionnelle est facile à mettre en œuvre et possède une sensibilité nettement améliorée par rapport au spectromètre à onde continue. Lors de l'impulsion à une fréquence proche de la fréquence de résonance des protons de la molécule, toutes les aimantations des protons vont basculer d'un angle de 90° . Le retour à l'équilibre, effectué en présence du champ magnétique, en précession libre est alors enregistré. Nous obtenons un signal dit de Free Induction Decay (FID). Le traitement mathématique de ce signal permet d'obtenir le spectre RMN. La réponse de chaque atome possédant un spin non nul, appelée résonance, se traduit par un pic plus ou moins complexe. On appelle déplacement chimique la variation de la fréquence de résonance d'un atome dans un environnement donné. Les atomes à spin non nul engendrent eux aussi des champs magnétiques qui sont détectables sur la résonance des atomes voisins par le phénomène de couplage. Une molécule peut être caractérisée par l'ensemble des déplacements chimiques et des couplages des atomes qui la composent.

Lors de la synthèse du naphthalène, du ^{12}C -phénanthrène et du ^{13}C -phénanthrène, les intermédiaires de chaque étape ont été analysés par RMN du proton (^1H) et du carbone (^{13}C). Il faut noter que dans le cas de la molécule ^{12}C , on observe les déplacements chimiques du carbone ^{13}C qui représente naturellement 1% du carbone des molécules non marquées. Dans le produit marqué, les couplages ^{13}C - ^{13}C et ^{13}C - ^1H compliquent considérablement le spectre.

Expérimentalement, 10 mg de produit sec sont dissous dans 2 ml de chloroforme deutéré (CDCl_3). Les spectres RMN (^1H et ^{13}C) sont enregistrés à température ambiante à partir d'un spectromètre Bruker Avance 300 MHz équipé d'une sonde multi-noyaux QNP. Le solvant deutéré est utilisé comme référence. Son pic résiduel en RMN du proton (^1H) se trouve à 7,26 ppm. Les signaux FID (Free Induction Decay) sont traités par le logiciel Topspin.

1.2.2 Analyse de l'acide 3-(2-naphtoyl) propionique (2)

Le produit de la réaction est composé de deux isomères (Figure IV.3) puisque l'anhydride succinique réagit sur deux carbones. Le déplacement chimique des protons impliqués dans un cycle aromatique génère des signaux entre 6 et 8,5 ppm. Le déplacement chimique des protons impliqués dans des chaînes aliphatiques de type CH_2 se traduit généralement par des signaux entre 3 et 4,5 ppm (Roussel, 1996)

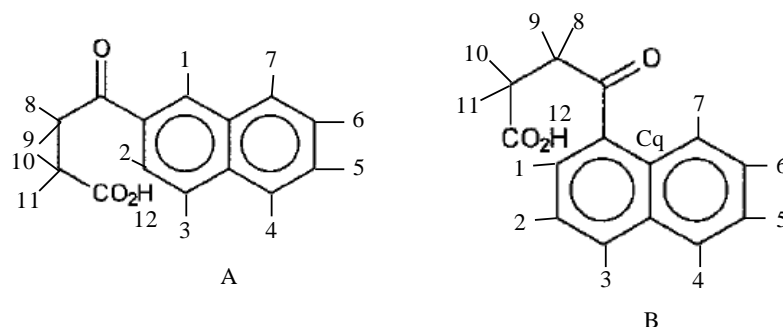


Figure IV.3 : Molécules annotées du produit de la condensation de l'anhydride succinique sur le naphthalène.

Au niveau des noyaux aromatiques, les protons A1 et B1 sont chimiquement différents dans les deux isomères. Le proton A1 est entouré de carbones quaternaires et est éloigné de l'influence des autres protons. Étant donné que la majorité des couplages observables sont en $3J$ (3 liaisons entre les atomes), il sera peu sensible aux autres protons de la molécule et apparaîtra sous forme d'un singulet. Le proton B1 n'a qu'un seul proton (B2) dans son environnement chimique. Les autres protons n'affectent que très peu son déplacement chimique, il formera un doublet. Les autres protons des noyaux aromatiques sont équivalents dans les deux isomères et formeront des multiplets dont les déplacements chimiques se superposeront au niveau du spectre. L'insert « Aromatique » du spectre Figure IV.4 confirme cette prédiction, bien que les multiplets superposés forment des massifs de signaux mal résolus.

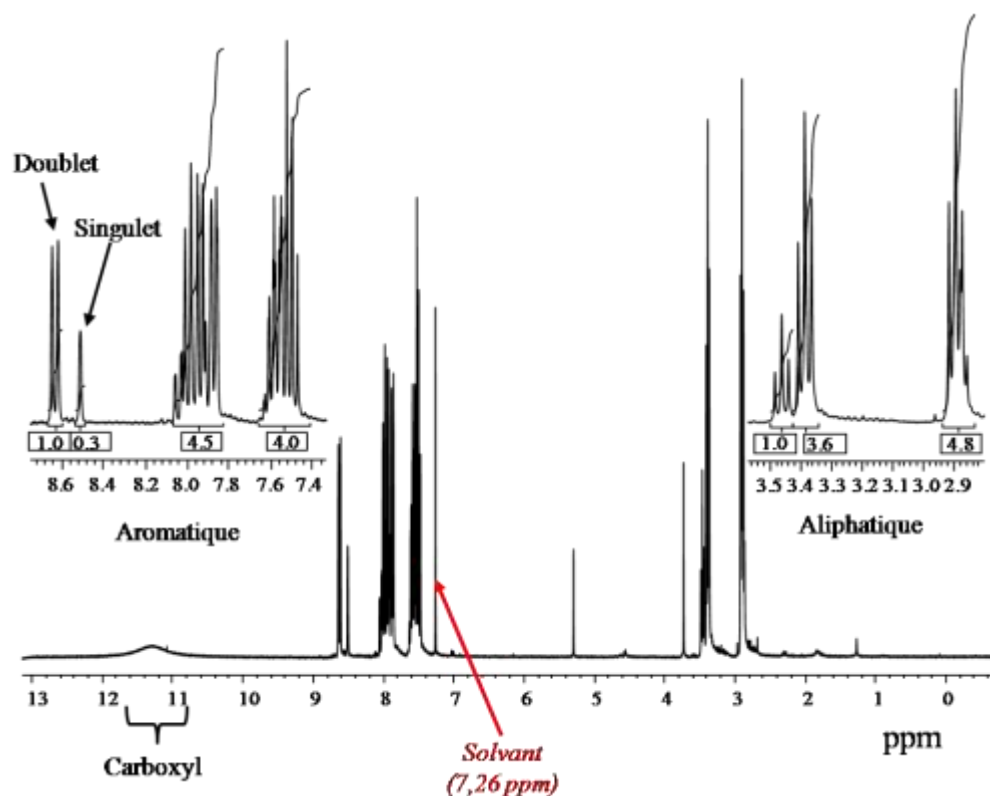


Figure IV.4 : Spectre RMN ^1H de l'acide 3-(2-naphtoyl) propionique (2). Les déplacements chimiques sont exprimés en ppm. Les valeurs représentées en dessous des spectres zoomés représentent l'intégration des aires sous pic.

Le proton du groupement carboxylique (COOH) de la chaîne aliphatique sort sous forme d'un signal large à 11,25 ppm, ce qui traduit le caractère échangeable de ce dernier.

Les deux positions de branchement du groupement C=O sur le cycle aromatique correspondant aux 2 isomères apparaissent clairement sur le spectre. Les protons les plus proches de ces positions ont des déplacements chimiques différents en fonction de l'isomère considéré (A ou B). Leur environnement étant peu différent, les signaux des protons aliphatiques A8 et A9 sortent donc à des valeurs de déplacements chimiques différents mais assez proches des signaux des protons aliphatiques B8 et B9. De plus, dans les deux isomères chaque proton 8 et 9 est influencé par les protons 10 et 11. Leur spectre produit des triplets bien distincts représentés à gauche dans l'insert « Aliphatique » (Figure IV.4)

Enfin, les protons A10, A11 et B10, B11, présentent un environnement chimique similaire avec un groupement carboxylique d'un côté et deux protons aliphatiques de l'autre. Les

triplets de leur spectre RMN sont superposés et forment un multiplet à droite de l'insert aliphatique.

Le spectre de la Figure IV.4 confirme que les deux isomères attendus sont présents dans le mélange réactionnel.

L'intégration de chaque massif est proportionnelle au nombre de protons impliqués. En présence des deux isomères en quantité égale, le singulet et le doublet des aromatiques devraient avoir une intégrale égale puisqu'ils représentent chacun 1 proton. Si on fixe la valeur de l'intégrale du doublet de l'isomère B à 1, l'intégrale du singulet de l'isomère A est calculée comparativement à cette valeur. Le ratio des aires permet alors de donner la proportion de chaque isomère. Ainsi, on obtient 23% (0,3/1,3) de A et 77% (1/1,3) de B. De même, les deux triplets de gauche des aliphatiques appartiennent chacun à l'un des isomères. Le rapport des aires de triplet donne 22% (1/4,6) d'un isomère et 78% (3,6/4,6) de l'autre.

Compte tenu des erreurs d'intégration, on peut conclure que la synthèse a produit les deux isomères de **(2)** dans les proportions suivantes : 23% de A et 77% de B.

Au cours des étapes de synthèse, la même analyse a été réalisée pour chaque intermédiaire afin de s'assurer de l'efficacité de la réaction.

1.2.3 Spectres RMN du phénanthrène

Les spectres des molécules marquées sont plus complexes à analyser. En effet, la présence de ^{13}C en forte proportion dans les molécules intensifie les couplages carbone-proton et carbone-carbone. Afin de s'assurer de la réussite des étapes de synthèse, les massifs dans les spectres RMN des produits ^{13}C ont été comparés aux pics dans les spectres des produits ^{12}C .

Pour plus de clarté, ne seront présentés ici que les spectres du phénanthrène.

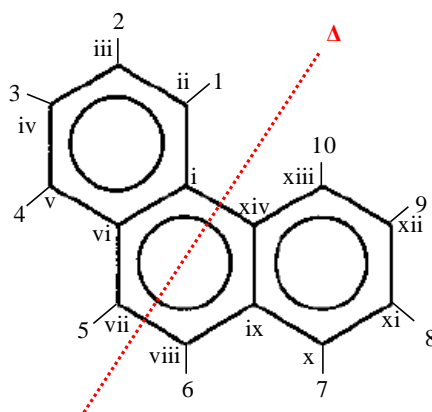


Figure IV.5 : Phénanthrène annoté. Les protons sont notés de 1 à 10 et les carbones de i à xiv. La droite Δ rouge représente l'axe de symétrie de la molécule.

La molécule de phénanthrène est symétrique par rapport à un axe Δ (Figure IV.5). Ainsi, seule la description de la moitié de la molécule est nécessaire. Par exemple, les protons H1 et H10 sont symétriques et seront influencés par un proton immédiatement voisin (H2 et H9 respectivement). Ils sont détectés sur le spectre par un doublet. Il en va de même pour les protons H4 et H7 qui sont représentés par un second doublet. Les protons H5 et H6 ne sont pas influencés directement par un autre proton, leur déplacement chimique apparaît sous la forme d'un singulet. Enfin, les protons H2, H3 (ou H8, H9) ont dans leur environnement chacun deux autres protons. Leur déplacement chimique apparaît sous la forme de triplets. H2 et H3 ayant un environnement chimique très proche, les deux triplets sont superposés, formant ainsi un multiplet. Ces quatre motifs sont retrouvés dans le spectre du ^{12}C -phénanthrène synthétisé au laboratoire (Figure IV.6).

Le spectre du ^{13}C -phénanthrène a été réalisé dans les mêmes conditions avec le déplacement chimique spécifique du chloroforme deutéré comme référence. Les motifs précédemment décrits ne sont plus visibles. Ils ont été remplacés par des pics plus larges et nettement moins bien résolus. Bien que les déplacements chimiques aient aussi variés sous l'influence du ^{13}C , la succession « doublet - doublet - multiplet - singulet - multiplet » peut être détectée (Figure IV.6). Afin de confirmer que la molécule obtenue est celle attendue, un spectre RMN du proton découplé du carbone a été réalisé. Dans cette expérience, on supprime artificiellement les couplages C-H pour ne laisser apparaître que les H-H. Cela revient finalement à considérer la molécule sans ^{13}C . Le résultat obtenu confirme cela puisque le spectre correspond parfaitement à celui du phénanthrène ^{12}C préparé indépendamment.

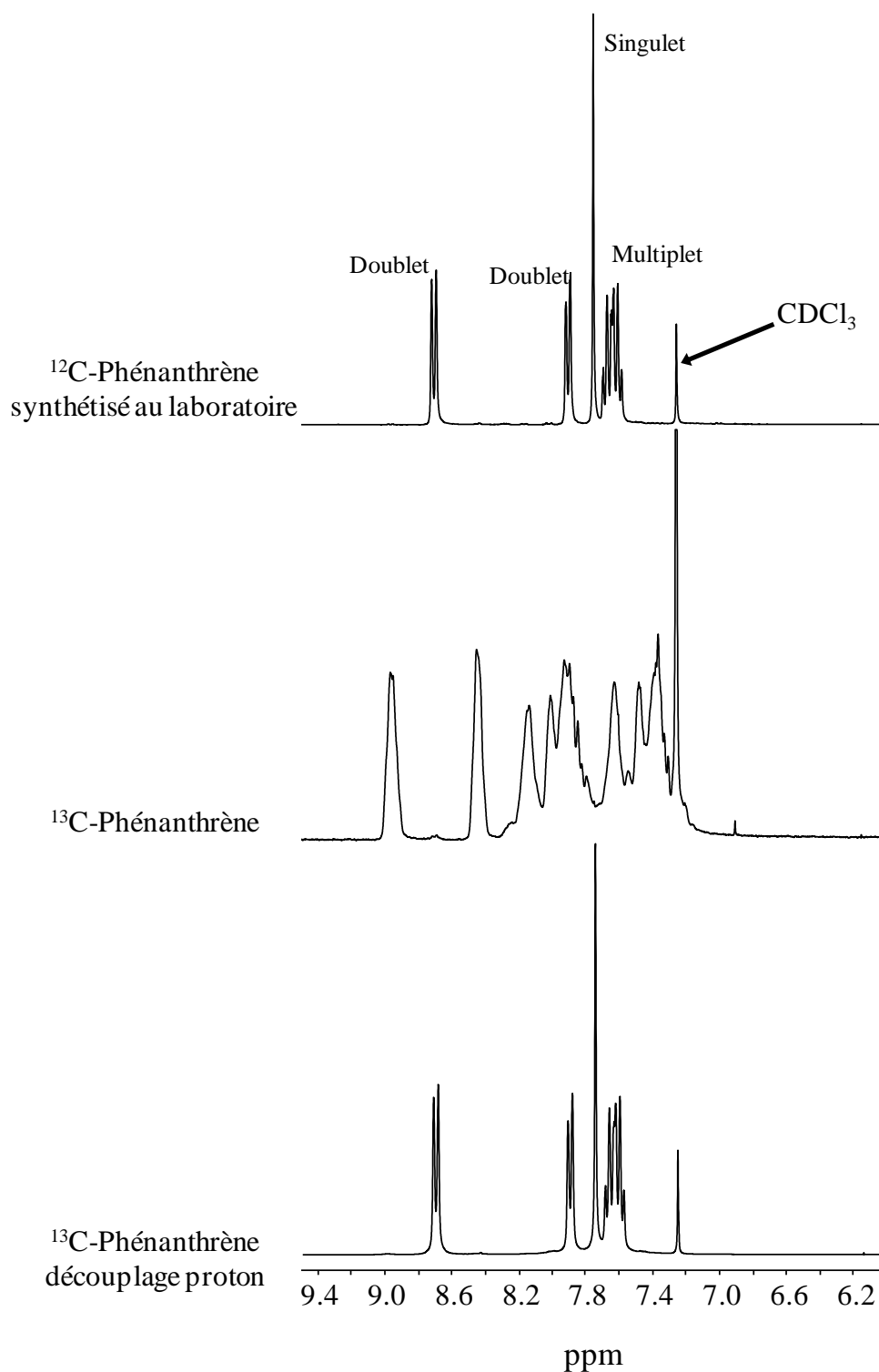


Figure IV.6 : Spectre RMN ¹H du phénanthrène. Les trois spectres ont été réalisés dans le chloroforme deutéré (CDCl₃). Ils ont été alignés sur la base du pic caractéristique de ce solvant à 7,26 ppm.

De la même manière, le spectre ^{13}C -RMN du phénanthrène montre deux pics q1 et q2 correspondant aux quatre carbones aromatiques quaternaires symétriques deux à deux i - iv et ix - xiv. Les 10 carbones aromatiques restant sont symétriques deux à deux et présentent un environnement chimique qui leur est propre. Ils devraient donner cinq pics distincts à des déplacements chimiques compris entre 110 et 175 ppm. Sur la Figure IV.7, seulement 4 pics aromatiques apparaissent.

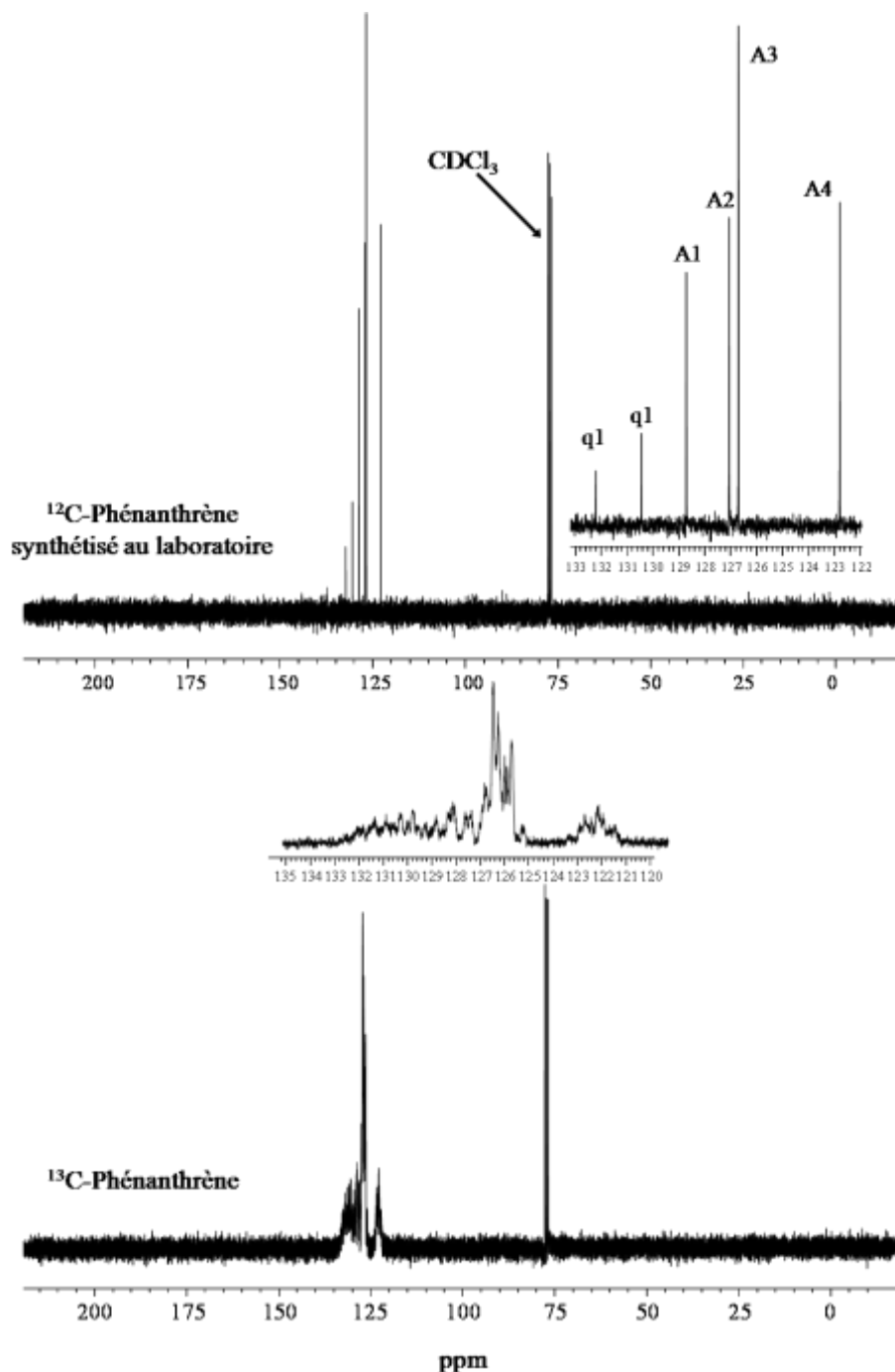


Figure IV.7 : Spectre RMN ^{13}C du phénanthrène. Les deux spectres ont été réalisés dans le chloroforme deutéré (CDCl_3). Ils ont été alignés sur la base du pic caractéristique de ce solvant à 76 ppm. Les carbones notés Ax portent un proton alors que les carbones quaternaires sont notés qx.

Si l'on fait abstraction de la multiplicité des signaux pour le spectre du produit marqué, on constate que les déplacements chimiques des deux produits correspondent bien. D'une manière générale, les signaux des carbones secondaires sont plus intenses que ceux des quaternaires (problème de relaxation), ce qui permet une attribution aisée. On note que les carbones xi et xii ont un environnement chimique suffisamment proche pour avoir le même déplacement chimique. Ils sont représentés sous le pic A3 dont la hauteur est le double des autres pics des carbones aromatiques.

Les spectres RMN ^1H et ^{13}C montrent que le produit qui a été synthétisé est bien le phénanthrène. A titre de comparaison, l'anthracène, qui est l'HAP de trois cycle linéaire possède deux axes de symétrie. Son spectre proton serait beaucoup plus simple et ne présenterait que deux doublets et un singulet. En RMN ^{13}C , quatre signaux devraient être observés, à savoir trois carbones aromatiques et un carbone quaternaire.

1.3 Chromatographie en phase Gazeuse et Spectrométrie de Masse (GC/MS)

Le marquage uniforme du phénanthrène au ^{13}C a été vérifié par chromatographie gazeuse et spectrométrie de masse (GC/MS ; $m/z = 192$ (M^+)) suivant le protocole décrit dans matériel et méthodes (paragraphe III.3.1.1).

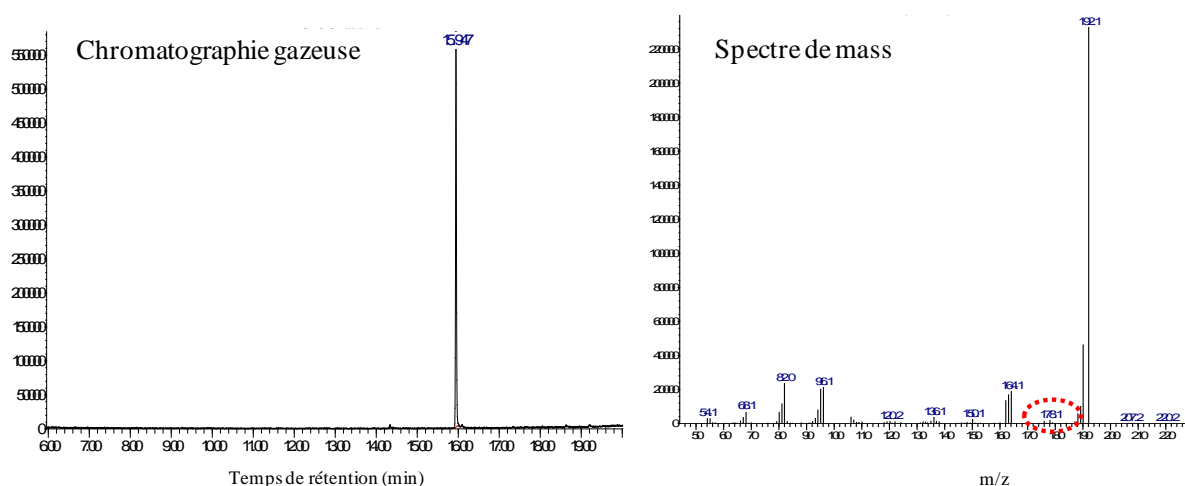


Figure IV.8 : Spectre CG/MS du ^{13}C -phénanthrène : temps de rétention théorique 16 min, signature $m/z = 192$. La signature m/z du ^{12}C -phénanthrène (178) est entourée en rouge.

Le temps de rétention du composé synthétisé (15,947 min) est le même que celui du ^{12}C -phénanthrène commercial (15,952). Le ^{13}C -phénanthrène a une masse (m/z) de 192 (M^+) ce qui correspond à la masse attendue (Figure IV.8). Le ^{12}C -phénanthrène ($m/z = 178$ (M^+)) est indétectable dans notre préparation.

Le composé synthétisé est donc bien le phénanthrène uniformément marqué avec l'isotope ^{13}C du carbone. Il sera utilisé comme substrat marqué dans les expériences SIP (Chapitre II et Chapitre III).

2. Dosage des HAP dans le sol

Les échantillons de sol qui ont été utilisés pour cette étude proviennent de lits à macrophytes, des bassins plantés de roseaux, utilisés pour le traitement des eaux de ruissellement de chaussées. Cette installation est située sur le bord de l'autoroute A43 Nord près de Chambéry (Savoie). Un premier bassin de rétention, recevant les eaux brutes provenant de l'autoroute, est constitué essentiellement d'une couche de matières organiques issues de la décomposition des végétaux reposant sur une couche d'argile. Après un déshuileur équipé d'une vanne, un bassin de traitement composé d'un lit de graviers, a été conçu pour éliminer les polluants organiques et retenir les polluants métalliques (Figure III.1 dans Matériel et Méthodes)

Les prélèvements ont été effectués sur une profondeur de 10 cm à partir de la surface à l'entrée de chacun des bassins.

Les sols ont été tamisés afin d'éliminer les graviers et les fragments de roseaux. Il a été déterminé, par pesée des masses humides et sèches que ces sols présentaient 50-60% d'humidité.

Les 16 HAP désignés prioritaires par l'EPA (Environmental Protection Agency, Etats-Unis) ont été dosés par GC/MS après extraction du sol en solvant comme décrit dans le chapitre Matériels et Méthodes.

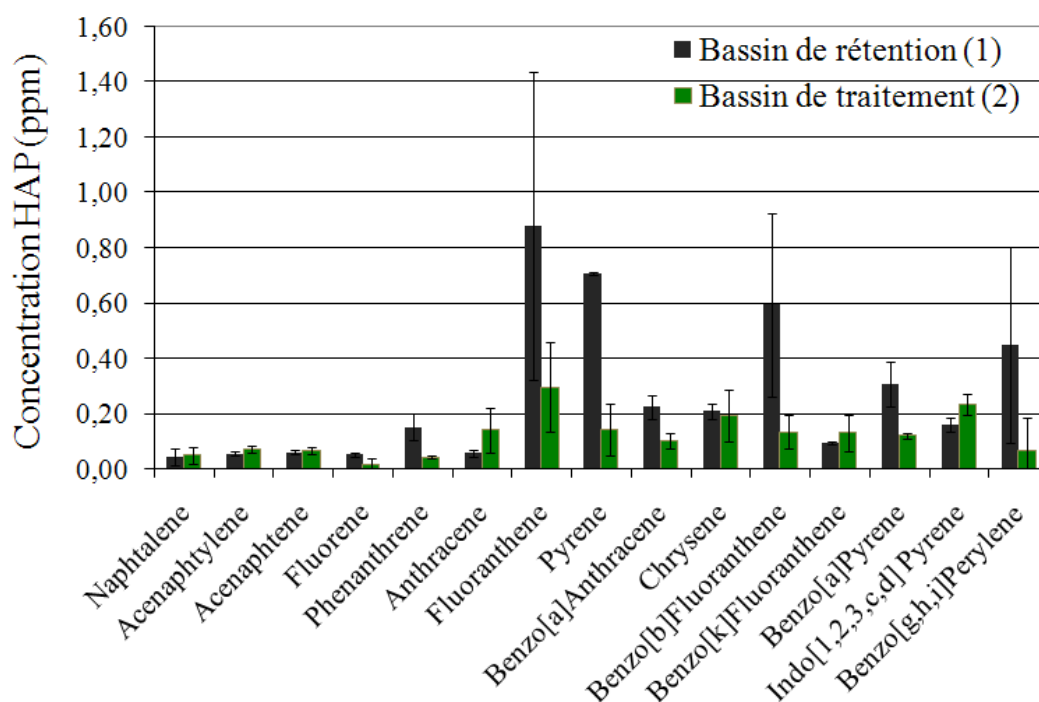


Figure IV.9 : Concentrations des 16 HAP prioritaires dans les boues des deux bassins de traitement des eaux de chaussée. Chaque valeur représente la moyenne de trois répliquats d'extraction, les barres d'erreur représentant l'écart type. Les concentrations sont exprimées en ppm (mg d'HAP/kg de sol sec)

Comme on pouvait s'y attendre, le bassin de rétention recevant les eaux brutes est deux fois plus concentré en HAP ($4,05 \pm 0,1$ ppm) que le bassin de traitement situé en aval ($1,80 \pm 0,06$ ppm). Le fluoranthène, le pyrène et le benzo[b]fluoranthène sont les HAP les plus concentrés dans le bassin de rétention (Figure IV.9). La concentration du phénanthrène dans ce même bassin ($0,15 \pm 0,05$ ppm) est 3,5 fois plus élevée que dans le bassin de traitement ($0,043 \pm 0,007$ ppm).

Les boues du bassin de rétention (1) sont donc plus chargées en HAP et en matières organiques que celles du bassin de traitement (2), tout au moins au niveau de la couche de graviers. De plus, les graviers étaient enrobés d'une faible quantité de matière organique, insuffisante pour la mise en place des microcosmes à partir des échantillons de ce bassin (2). Le bassin de rétention, susceptible de contenir une population bactérienne plus dense et plus exposée aux HAP, a été jugé le plus propice à l'étude de biodiversité par l'approche SIP.

3. Etude préliminaire en microcosmes exposés au ^{12}C -phénanthrène

Cette étude préliminaire avait pour objectifs (1) de s'assurer que la dose de phénanthrène ajoutée n'affectait pas de manière négative l'activité microbienne du sol, (2) de déterminer la quantité de substrat adéquate pour suivre sa dégradation, (3) de définir le temps d'incubation nécessaire pour cibler les bactéries actives tout en évitant le phénomène de cross-feeding.

3.1 Préparation des microcosmes

Les microcosmes ont été préparés à partir de 50 g de sol humide du bassin de rétention auxquels ont été ajouté du ^{12}C -phénanthrène à des concentrations 53 fois (8 ppm ; Ex50) ou 330 fois (50 ppm ; Ex300) supérieures à la teneur initiale du sol. Un essai contrôle, sans ajout de phénanthrène et un essai abiotique ont aussi été préparés. Ils ont été incubés 22 jours à 25°C, à l'abri de la lumière et sans agitation.

3.2 Cinétiques comparées de minéralisation dans les microcosmes

La matière organique est transformée par les microorganismes pour partie en CO_2 et pour partie en composés carbonés qui sont incorporés dans les voies de biosynthèse. Le suivi de la minéralisation de la matière organique et du phénanthrène dans les microcosmes a été réalisé par le dosage du CO_2 par GC/MS.

La teneur en CO_2 mesurée dans les microcosmes abiotiques a été retranchée à la production de CO_2 dans les autres essais.

Après un temps de latence très court (12 h pour Ex50 et Ex300, imperceptible pour le Contrôle), la production de CO_2 dans tous les microcosmes a d'abord suivi une cinétique rapide pendant 4 à 7 jours, puis atteint un plateau après environ 10 jours sauf pour le contrôle où une cinétique plus lente s'est établie (Figure IV.10). Le ralentissement de production de CO_2 pourrait s'expliquer par une diminution de la pression partielle en oxygène au niveau du sol, du fait de l'absence d'agitation des microcosmes.

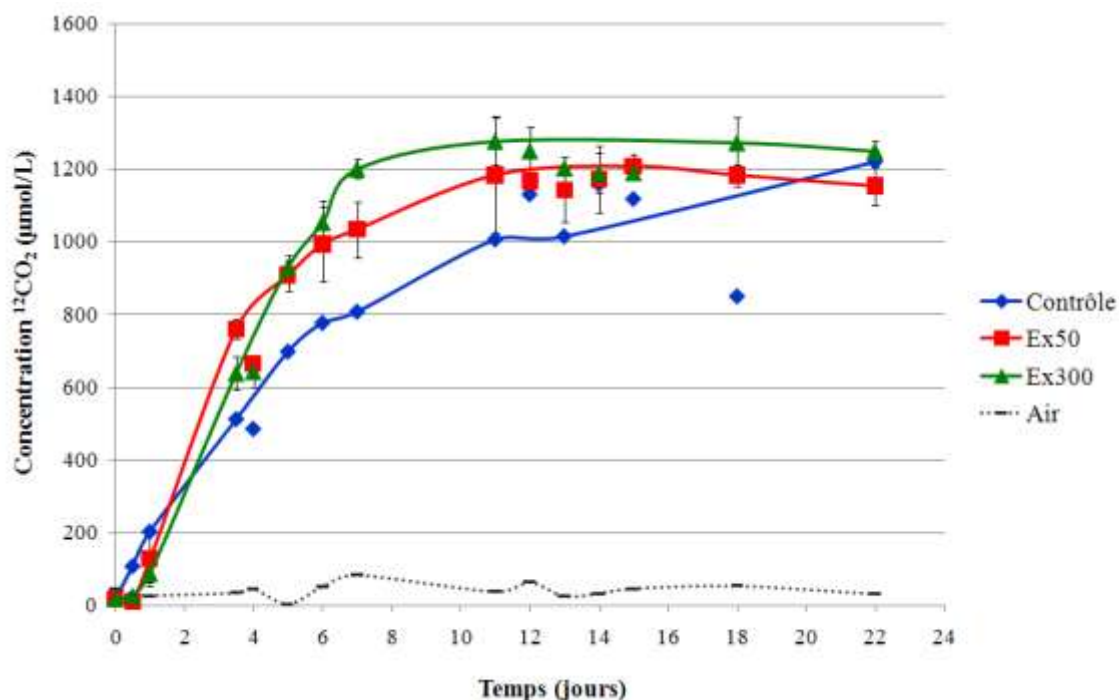


Figure IV.10 : cinétiques de minéralisation dans les microcosmes incubés avec ou sans phénanthrène. Les valeurs de concentration dans les microcosmes avec phénanthrène sont des moyennes et écarts types de duplicats. L'incubation sans phénanthrène (Contrôle) a été réalisée en un seul exemplaire. Les concentrations sont exprimées en μmol de CO_2 par litre de phase gazeuse des microcosmes.

La respiration dans les microcosmes supplémentés avec l'hydrocarbure n'a pas été inhibée par rapport à celle du microcosme de contrôle ce qui suggère que les doses de phénanthrène ajoutées n'ont pas eu d'effet toxique sur les populations bactériennes du sol étudié.

Au contraire, entre 5 et 7 jours, la quantité de CO_2 produite dans les microcosmes avec phénanthrène était supérieure à celle mesurée dans le microcosme contrôle. Cette observation suggère que l'augmentation transitoire du métabolisme global résulte de la minéralisation du phénanthrène. Cependant, on ne peut accorder qu'une confiance relative à ce résultat dans la mesure où l'expérience en microcosme contrôle n'a été réalisée qu'une seule fois.

Le taux maximal de production de CO_2 mesuré dans les microcosmes enrichis en phénanthrène est supérieur (Ex50 : $250 \mu\text{M/j}$; Ex300 : $181 \mu\text{M/j}$) à celui mesuré dans le microcosme contrôle ($122 \mu\text{M/j}$). Toutefois, la différence est assez faible ce qui indique que l'essentiel du CO_2 produit par les communautés bactériennes provient de la minéralisation de la matière organique du sol. Le phénanthrène étant un assez mauvais substrat carboné, il ne stimulerait que légèrement la respiration dans le microcosme. On pourrait s'attendre à ce que le taux mesuré dans les microcosmes Ex 300 soit supérieur à celui des microcosmes Ex50. Or

c'est l'inverse qui se produit, ce qui pourrait être la conséquence d'une inhibition partielle du métabolisme à la plus forte dose de phénanthrène.

On peut toutefois constater que, quelle que soit la quantité de substrat ajouté aux microcosmes, il est difficile de distinguer clairement la part du métabolisme imputable au substrat ajouté. Dans les expériences de marquage SIP, il conviendra d'ajouter une quantité de phénanthrène marqué au moins équivalente à la plus forte dose testée afin que la production de CO₂ soit observable au dessus du bruit de fond. D'autre part, comme la minéralisation atteint un palier à 10 jours, une durée d'incubation inférieure à 15 jours devrait suffire à marquer les microorganismes capables de métaboliser le phénanthrène.

3.3 Estimation de la diversité bactérienne dans les sols incubés avec le phénanthrène

Cette expérience préliminaire de contamination des microcosmes par du phénanthrène a été l'occasion d'estimer grossièrement la diversité bactérienne du sol étudié. Des échantillons de 5 g de sol ont été prélevés avant (T0) et à la fin de l'incubation (Tf) pour en extraire l'ADN et analyser la richesse en séquences d'ARNr 16S.

3.3.1 Amplification et clonage des gènes d'ARNr 16S, marqueurs de l'identité bactérienne

L'ADN génomique de ces échantillons a été extrait à l'aide du kit UltraClean™ Mega Soil DNA Kit (MoBio) selon les instructions du fabricant. L'ADN obtenu est de bonne qualité selon l'analyse réalisée sur gel d'agarose. Le rendement calculé, 276 µg d'ADN par g de sol sec en moyenne, est très probablement surestimé car les préparations sont légèrement contaminées par une substance de couleur jaune.

Les gènes d'ARN ribosomaux 16S (ARNr 16S) ont été amplifiés par PCR à l'aide des amorces universelles 27f/1492r (Annexe 1) et de l'enzyme KOD comme décrit dans Matériels et Méthode. Les produits PCR bouts francs de 1465 pb ont été clonés à l'aide du kit Perfectly Blunt®. Entre 200 et 500 clones ont été obtenus pour les essais Ex50 et Ex300, ainsi que pour

l'expérience contrôle. Dix clones de chaque expérience ont été criblés par PCR utilisant les amorces 27f/1492r afin de vérifier l'insertion du gène dans le vecteur.

Parallèlement, le même protocole a été appliqué à de l'ADN génomique de *Sphingomonas* sp. CHY-1 et de *Mycobacterium* sp. 6PY1, des souches bactériennes étudiées au laboratoire pour leur faculté de dégrader les HAP. Leur gène d'ARNr 16S a été amplifié et cloné et cinq clones par essai ont été criblés. Enfin, un témoin de clonage a été réalisé dans lequel aucun insert n'a été ajouté.

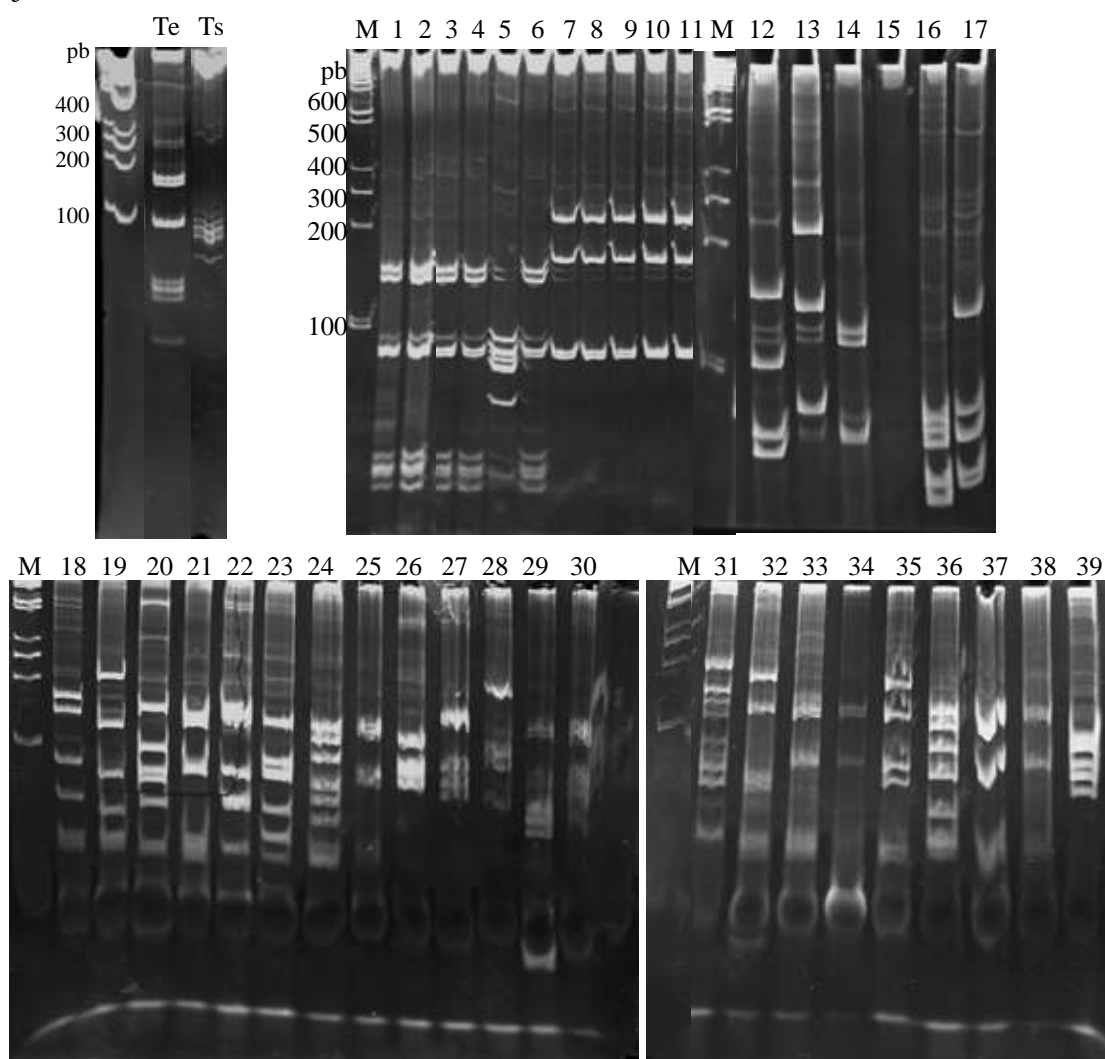


Figure IV.11 : Profils RFLP des gènes d'ARNr16S issus des microcosmes incubés avec ^{12}C -phénanthrène.

M : marqueur de taille 100 pb (Fermentas)

Te : Gène d'ARNr 16S de *Escherichia coli*

Ts : Gène d'ARNr 16S de *Sphingomonas* sp. CHY-1

Puits 1 et 2 : témoins de clonage

Puits 3 à 6 : Clones d'ARNr 16S de *Sphingomonas* sp. CHY-1

Puits 7 à 11 : Clones d'ARNr 16S de *Mycobacterium* sp. 6PY1

Puits 12 à 21 : Gènes d'ARNr16S de l'échantillon Contrôle

Puits 22 à 30 : Gènes d'ARNr16S de l'échantillon Ex50

Puits 31 à 39 : Gènes d'ARNr16S de l'échantillon Ex300.

3.3.2 Profils de polymorphisme de restriction (RFLP)

Les produits d'amplification ont été digérés par l'enzyme *AluI*, et les profils RFLP ont été analysés par électrophorèse sur gel de polyacrylamide 10% (Figure IV.11).

Les témoins de clonage (puits 1 et 2) produisent des profils identiques à ceux de *E.coli*. Le gène d'ARNr16S amplifié à l'aide des amorces universelles provient de l'ADN génomique de la souche de clonage. Tous les clones présentant ce profil seront considérés comme des clones négatifs.

Le clonage du gène d'ARNr 16S de *Sphingomonas* sp. CHY-1 ne produit le profil attendu que pour 1 clone sur 4. Le clonage du gène d'ARNr 16S de *Mycobacterium* sp. 6PY1 produit 5/5 profils identiques au profil attendu. L'efficacité de clonage est donc variable.

Pour le clonage des gènes d'ARNr 16S issus des microcosmes, j'ai obtenu 3 clones négatifs sur 10 (puits 14, 15 et 21) pour l'essai Contrôle, 1 clone négatif sur 9 (puits 25) pour Ex50 et 4 clones négatifs sur 9 (puits 33, 34, 37 et 38) pour Ex300.

La diversité des profils obtenus dans ces trois essais est relativement élevée. Les gènes d'ARNr16S amplifiés à partir de l'ADN Contrôle, Ex50 et Ex300 présentent 7/7, 7/8 et 4/5 profils différents.

Ce résultat est logique et suggère que le sol étudié recèle une microflore bactérienne variée, même si le nombre de clones analysés était insuffisant du point de vue statistique.

4. Conclusions

Cette étude préliminaire a mis en évidence la nécessité d'apporter le phénanthrène en quantité suffisante, au minimum 200 à 300 fois la concentration initiale dans le sol, afin de s'assurer que le métabolisme de ce substrat soit détectable par dessus le bruit de fond lié à la minéralisation de la matière organique du sol. De plus, la cinétique de production de CO₂ montre qu'une incubation de quinze jours est suffisante pour marquer les bactéries qui dégradent le phénanthrène. Enfin, l'estimation de la diversité basée sur une analyse succincte de profils d'ARNr16S dans l'ADN du sol laisse présager d'une grande richesse génotypique. En revanche, la méthode de criblage par RFLP n'est pas adaptée à une caractérisation à grande échelle des populations bactériennes impliquées dans la dégradation des HAP.

Chapitre II. Réponse des populations bactériennes du sol à un épisode de contamination au phénanthrène

Dans le premier chapitre, nous avons vu que l'ajout de phénanthrène dans le sol avait peu d'incidence sur l'activité métabolique aérobie des communautés bactériennes présentes.

Dans ce qui suit, nous donnerons une description du système expérimental mis en place pour analyser, par marquage isotopique (SIP), la diversité bactérienne du sol associée à la dégradation des HAP. Avant d'aborder l'analyse phylogénétique de cette diversité, qui sera présentée dans le Chapitre III, nous avons cherché à connaître l'impact global d'un HAP, ajouté à deux concentrations différentes, sur la communauté bactérienne du sol.

Nous nous sommes donc demandé (1) si la composition globale des communautés bactériennes du sol changeait suite à un pic de pollution par un HAP, (2) dans quelle mesure et sur quelle échelle de temps certaines bactéries du sol étaient capables de dégrader ce HAP. En essayant d'apporter des réponses à ces questions, nous souhaitons obtenir des éléments pour juger de l'efficacité de la bioremédiation pour traiter la pollution aux hydrocarbures générée par les autoroutes.

Afin de mettre en évidence cette activité métabolique spécifique, nous avons utilisé un HAP modèle, le phénanthrène, uniformément marqué au ^{13}C (^{13}C -Phe). Sa minéralisation conduit à la production de $^{13}\text{CO}_2$ qui est quantifiable indépendamment du CO_2 naturel par GC/MS.

1. Cinétique de la minéralisation du ^{13}C -phénanthrène en microcosmes

1.1 Description des microcosmes

Pour les besoins des expériences de marquage SIP décrites dans ce chapitre et le suivant, une série de microcosmes contenant chacun 20 g de sol a été préparée. Dans certains d'entre eux, ^{13}C -Phe a été ajouté à une concentration de 33 ppm ou bien de 337 ppm, c'est-à-dire 200 fois (SIP200) ou 2000 fois (SIP2000) supérieure à la teneur initiale du sol en phénanthrène. Des microcosmes de contrôle, sans ajout de phénanthrène (Contrôle négatif) et en conditions abiotiques ont aussi été préparés, ainsi que des microcosmes ayant reçu une dose de phénanthrène non marqué 200 fois supérieure à la teneur initiale (33 ppm ; Contrôle positif). Enfin, un microcosme ayant reçu la dose maximale de ^{13}C -Phe a été inoculé avec

Sphingomonas sp. CHY-1 connu pour sa capacité à dégrader le phénanthrène (Demaneche *et al.*, 2004) (SIP2000 CHY1).

L'incubation a été réalisée à 25°C pendant 14 jours à l'obscurité. Après 11 jours, les microcosmes ont été ouverts 30 min afin de rééquilibrer la teneur en O₂.

1.1 Impact d'une dose élevée de phénanthrène sur l'activité métabolique des bactéries du sol

Les populations bactériennes aérobies se développent en métabolisant des substrats carbonés du sol, produisant ainsi du ¹²CO₂. L'ajout d'une quantité élevée de HAP potentiellement toxique pourrait se traduire par une diminution de l'activité métabolique et donc de la production de CO₂. Nous avons testé cette hypothèse par suivi de la production du ¹²CO₂ dans les microcosmes.

Dans les microcosmes abiotiques, le dégagement de ¹²CO₂ reste à un niveau basal (1 mM soit 2,28% du contenu gazeux des microcosmes) tout au long de l'incubation (Figure IV.12). Il ne semble pas y avoir d'activité métabolique dans ces microcosmes. On peut supposer que le CO₂ détecté provient de l'émission d'une forme piégée dans le sol, du bicarbonate par exemple.

La production de ¹²CO₂ dans les autres microcosmes de l'expérience n'est pas précédée d'une phase de latence. Les microcosmes incubés sans ajout de phénanthrène présentent le taux de production de ¹²CO₂ le plus élevé (446 μM/j). Celui-ci diminue avec l'augmentation de la dose de phénanthrène ajoutée : Contrôle+ 371 μM/j, SIP200 357 μM/j, SIP2000 179 μM/j. Enfin, la production de ¹²CO₂ dans les microcosmes Contrôle négatif et SIP200 tend à ralentir après 8 jours d'incubation.

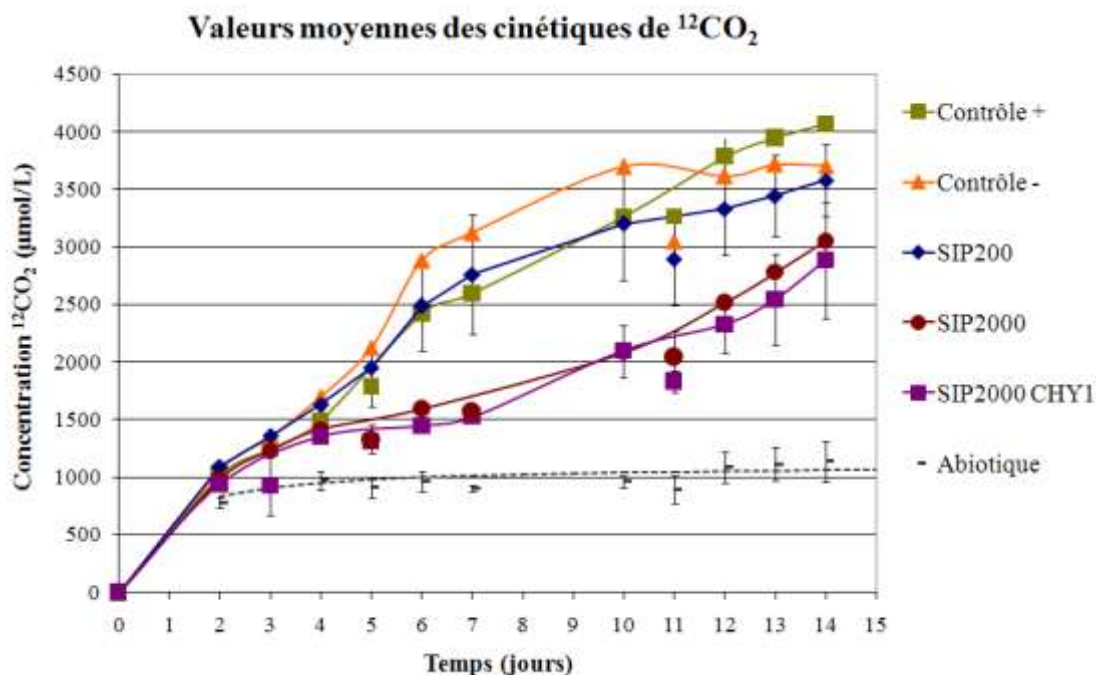


Figure IV.12 : Cinétiques de production de $^{12}\text{CO}_2$ dans les microcosmes de l'expérience SIP. Les concentrations ont été calculées à l'aide d'une gamme étalon de CO_2 . La teneur en $^{12}\text{CO}_2$ de l'air a été mesurée chaque jour et soustraite aux concentrations mesurées dans les flacons. Les microcosmes ont été ouverts après 11 jours d'incubation. Les moyennes et les écarts types des concentrations ont été calculés sur des répliquats d'incubation (SIP200 et SIP2000 : triplicats ; autres microcosmes : duplicats).

Ces constatations suggèrent que le phénanthrène a un faible impact sur l'activité métabolique des communautés bactériennes du sol tant que sa concentration initiale est inférieure ou égale à 33 ppm. A 337 ppm, une diminution de la respiration est observée entre 3 et 7 jours, puis l'activité reprend quasi normalement au-delà, sans doute parce que la majeure partie du phénanthrène a été consommé par les espèces spécialisées (Figure IV.14). A forte concentration, le phénanthrène provoque donc une inhibition transitoire de l'activité métabolique globale de la communauté microbienne.

1.2 Cinétique de transformation du ^{13}C -phénanthrène en $^{13}\text{CO}_2$

La forte dose de phénanthrène ajoutée dans les microcosmes SIP2000 inhibe la respiration des bactéries du sol, mais qu'en est-il de l'activité métabolique des populations capables de dégrader ce HAP ? Subissent-elles aussi une forme d'inhibition, ou sont-elles capables de supporter ces concentrations ?

Dans les microcosmes contrôles sans ajout de ^{13}C -Phe, le $^{13}\text{CO}_2$ généré provient de la respiration sur des substrats du sol contenant naturellement du ^{13}C (le carbone ^{13}C représente

1,1% du carbone naturel). Dans les microcosmes abiotiques, la quantité de $^{13}\text{CO}_2$ reste stable ($9,54 \pm 3,33 \mu\text{M}$ soit 0,95% du CO_2 total) au cours des 14 jours d'incubation (Figure IV.13) suggérant une nouvelle fois qu'il provient de la libération sous forme gazeuse du $^{13}\text{CO}_2$ piégé dans le sol.

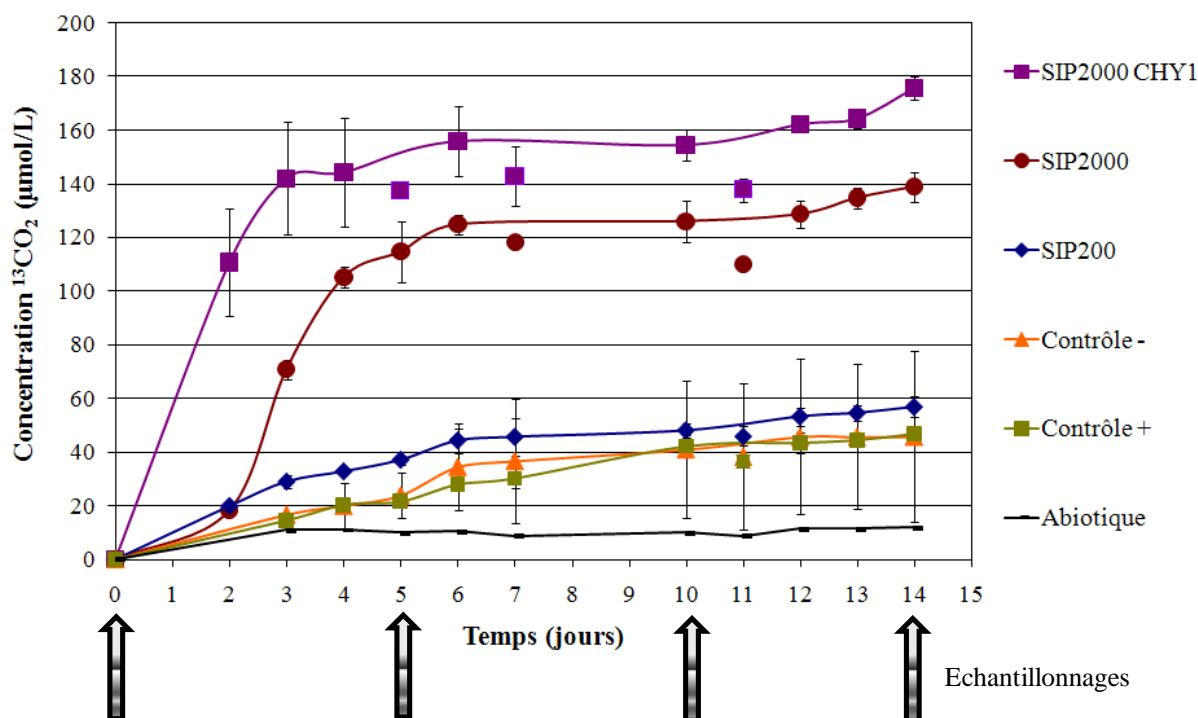


Figure IV.13 : Cinétiques de production de $^{13}\text{CO}_2$ dans les microcosmes de l'expérience SIP au cours des 14 jours d'incubation. Les microcosmes ont été ouverts après 11 jours. La teneur en $^{13}\text{CO}_2$ dans l'air a été mesurée chaque jour et soustraite aux concentrations mesurées dans les flacons. Les moyennes et les écarts types des concentrations ont été calculés à partir des valeurs dosées dans des réplicats d'incubation (SIP200 et SIP2000 : triplicats ; autres microcosmes : duplicats). Les flèches matérialisent les temps auxquels des échantillons ont été prélevés pour l'extraction d'ADN.

La production de $^{13}\text{CO}_2$ dans les microcosmes SIP200 est plus élevée que dans les contrôles sans substrat marqué entre 0 et 6 jours avec un taux de minéralisation de $7,39 \mu\text{M/j}$ (contre $4,67 \mu\text{M/j}$). En revanche, elle est à peine supérieure au bruit de fond détecté dans les microcosmes contrôles au-delà de 7 jours. Lorsque le composé marqué est apporté en plus faible quantité, sa minéralisation s'étend sur une période plus longue que lorsqu'il est fortement concentré.

Dans les microcosmes ayant reçu la dose maximale de $^{13}\text{C-Phe}$, une phase rapide de production de $^{13}\text{CO}_2$ est observée entre 2 et 4 jours, suivie par une phase plus lente. La production de $^{13}\text{CO}_2$ atteint un taux maximal ($43,4 \mu\text{M/j}$) après deux jours de latence puis se stabilise après cinq jours. Une cinétique similaire est observée dans les microcosmes SIP2000

CHY1, inoculés avec *Sphingomonas* sp. CHY-1 connue pour sa capacité à dégrader les HAP. Cependant, le taux maximal de minéralisation (47,3 $\mu\text{M/J}$) est atteint sans phase de latence. Ce résultat suggère que des microorganismes du sol aptes à minéraliser ^{13}C -Phe, à l'instar de *Sphingomonas* sp. CHY-1, se multiplient pendant le temps de latence observé dans les microcosmes SIP2000.

On peut donc s'attendre à ce qu'une population bactérienne spécifique se soit développée rapidement pour atteindre en 2 jours une taille critique suffisante pour minéraliser la quasi-totalité du ^{13}C -Phe. Cette hypothèse expliquerait le ralentissement brutal de la formation de $^{13}\text{CO}_2$ observé après 5 jours. Pour expliquer ce ralentissement, on peut avancer une autre hypothèse selon laquelle une partie du ^{13}C -Phe ajouté n'est pas directement accessible aux microorganismes parce qu'adsorbé ou piégé par les particules de sol.

Après ouverture des microcosmes (11 jours), la cinétique de production de $^{13}\text{CO}_2$ peut être interprétée comme la poursuite de la phase lente ou comme une reprise de la minéralisation favorisée par une augmentation de la pression partielle en oxygène au niveau du sol.

Pour essayer de mieux comprendre le devenir du phénanthrène dans les microcosmes, sa concentration résiduelle a été déterminée par deux méthodes : l'estimation indirecte de la quantité de substrat consommé et le dosage du ^{13}C -Phe restant dans les microcosmes.

La quantité théorique de substrat consommé a été calculée à partir des courbes de minéralisation. En déduisant la quantité de $^{13}\text{CO}_2$ produite dans les microcosmes sans ajout de ^{13}C -Phe (C_c) de celle produite dans les microcosmes SIP (C_s), on obtient la quantité nette de $^{13}\text{CO}_2$ associée à la minéralisation du ^{13}C -Phe. Si on rapporte cette quantité à la quantité théorique (Q_{th}) de CO_2 produite dans le cas où 100% de phénanthrène est minéralisé (1 molécule de phénanthrène donne 14 molécules de CO_2), la formule

$$\frac{(C_s - C_c) * Vf}{Q_{th}}$$

permet de calculer le pourcentage de ^{13}C -Phe transformé en $^{13}\text{CO}_2$ au cours de l'incubation dans un microcosme de volume V_f .

Dans les microcosmes SIP, 19% (SIP200), 16% (SIP2000) et 22% (SIP2000CHY-1) de ^{13}C -Phe ont été minéralisés après 14 jours. Or, la minéralisation d'une molécule de phénanthrène par voie aérobie conduit théoriquement à la production de 8 molécules de CO_2 (et non 14) soit environ 60% du substrat transformé en CO_2 . Ce calcul suggère donc qu'une partie seulement du ^{13}C -Phe a été consommée ou bien qu'une partie du CO_2 n'a pas été détectée parce qu'elle est restée piégée dans le sol. Le reste du carbone a probablement été transformé en biomasse ou bien en métabolites (voir paragraphe I.2.5).

En parallèle, le ^{13}C -Phe résiduel a été extrait et dosé dans l'expérience SIP2000 après 5 jours et 14 jours d'incubation (Figure IV.14). Les résultats obtenus montrent que plus de 90% du substrat a été consommé dans les cinq premiers jours de l'incubation. Si une telle proportion du phénanthrène avait été entièrement minéralisée, on devrait s'attendre à trouver dans la phase gazeuse une concentration en $^{13}\text{CO}_2$ au moins 2 fois plus élevée que celle trouvée expérimentalement.

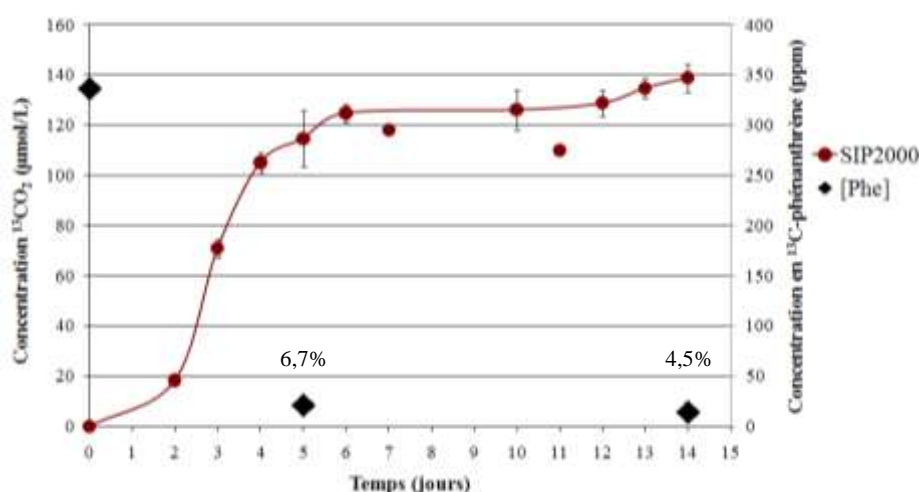


Figure IV.14 : courbe de minéralisation du ^{13}C -Phe dans les microcosmes SIP2000 et concentrations résiduelles de ^{13}C -Phe à 5 et 14 jours d'incubation.

La différence entre les deux estimations peut avoir trois explications :

- une partie du $^{13}\text{CO}_2$ produit a pu être piégée dans le sol sous forme de bicarbonates,
- une partie du $^{13}\text{CO}_2$ a pu être consommé par des bactéries autotrophes,
- une partie du ^{13}C -Phe est resté piégé dans le sol et donc n'a pas pu être extraite.

Malgré une incertitude sur la quantité exacte de substrat consommé, le ^{13}C -Phe a été dégradé en grande partie durant la phase rapide de la cinétique de minéralisation.

2. Effet du phénanthrène sur la composition globale de la communauté bactérienne du sol

Nous avons montré précédemment que l'ajout de phénanthrène à forte dose perturbe l'activité métabolique globale des microorganismes du sol. Cependant ce constat ne permet pas de savoir si la diversité bactérienne est affectée. La structure globale des communautés bactériennes a donc été examinée par deux approches moléculaires ciblant les gènes d'ARNr 16S : la comparaison des profils génotypiques basée sur des analyses SSCP et la quantification de phylums bactériens par qPCR.

Ces deux méthodes ont été mises en œuvre sur des préparations d'ADN métagénomique réalisées à 5, 10 et 14 jours au cours de l'expérience de marquage SIP (Tableau IV.1). Les extractions réalisées à l'aide de kits (MoBio) ont permis d'obtenir 6 µg d'ADN/g de sol à partir de 270 mg de sol et 25 µg d'ADN/g à partir de 10 g de sol (en masse humide).

La diversité bactérienne a été tout d'abord étudiée par comparaison des profils SSCP, en ciblant une région variable du gène d'ARNr 16S.

Tableau IV.1 : Récapitulatif des temps auxquels l'ADN a été extrait des microcosmes pour analyse

Temps (jours)	Microcosmes*	Masse (g)	Masse sèche (g)
0	Contrôle sans phénanthrène I	10,03	4,01
0	Contrôle sans phénanthrène II	10,07	4,03
0	Contrôle sans phénanthrène III	10,08	4,03
0	Contrôle sans phénanthrène IV	8,83	3,53
5	Contrôle avec phénanthrène I	11,50	4,03
5	Contrôle avec phénanthrène II	12,50	4,38
5	SIP200 I	10,13	3,55
5	SIP200 II	11,94	4,18
5	SIP200 III	11,74	4,11
5	SIP2000 I	11,76	4,12
5	SIP2000 II	10,59	3,71
5	SIP2000 III	12,02	4,21
5	SIP2000 CHY1 I	11,94	4,18
5	SIP2000 CHY1 II	12,14	4,25
10	Contrôle avec phénanthrène I	10,96	3,84
10	Contrôle avec phénanthrène II	11,80	4,13
10	SIP200 I	11,71	4,10
10	SIP200 II	11,59	4,06
10	SIP200 III	10,55	3,69
10	SIP2000 I	9,97	3,49
10	SIP2000 II	11,90	4,17
10	SIP2000 III	12,70	4,45
10	SIP2000 CHY1 I	12,29	4,30
10	SIP2000 CHY1 II	12,25	4,29
14	Abiotique I	11,93	4,18
14	Abiotique II	11,68	4,09
14	Contrôle sans phénanthrène I	11,30	3,96
14	Contrôle sans phénanthrène II	11,77	4,12
14	Contrôle avec phénanthrène I	10,70	3,75
14	Contrôle avec phénanthrène II	11,15	3,90
14	SIP200 I	11,23	3,93
14	SIP200 II	10,63	3,72
14	SIP200 III	10,68	3,74
14	SIP2000 I	9,65	3,38
14	SIP2000 II	11,28	3,95
14	SIP2000 III	11,59	4,06
14	SIP2000 CHY1 I	11,97	4,19
14	SIP2000 CHY1 II	11,97	4,19

*les chiffres romains représentent les répliqués d'échantillons

1.1 Comparaison des profils de diversité bactérienne dans les microcosmes incubés avec ou sans phénanthrène

La diversité bactérienne des microcosmes a été analysée par électrophorèse capillaire couplée à la SSCP (CE-SSCP). Une région variable du gène d'ARNr 16S, la région V3, de 200 nucléotides, a été amplifiée à l'aide d'amorces s'hybridant à des séquences conservées situées de part et d'autre. Les fragments obtenus ont été dénaturés pour générer des simples brins qui se replient en adoptant une conformation dépendant de leur séquence.

Les simples brins, rendus fluorescents par marquage au cours de la PCR, ont ensuite été séparés par électrophorèse capillaire. Le temps de migration des brins dans le capillaire dépend de leur conformation. On obtient un profil de fluorescence en fonction du temps, dans lequel les pics de fluorescence correspondent à un ou plusieurs fragments. L'abscisse est étalonnée en unités arbitraires (paires de bases), correspondant au temps de migration de marqueurs d'ADN double brin dans les mêmes conditions d'électrophorèse. Certains fragments simple brin ne donnent pas un seul pic mais 2 ou 3, comme c'est le cas du gène ribosomal de *Sphingomonas* CHY-1 (Figure IV.15).

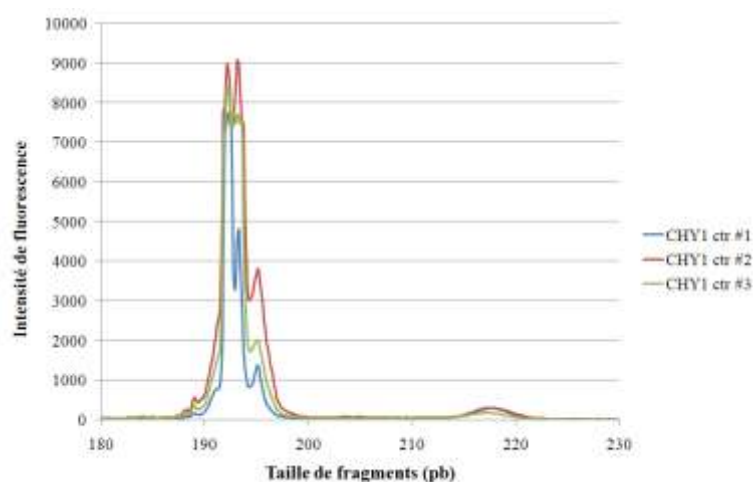


Figure IV.15 : Analyse SSCP de la région variable V3 du gène d'ARNr 16S de *Sphingomonas* sp. CHY-1. L'axe des abscisses est gradué en unités arbitraires correspondant au temps de migration dans les capillaires (Zinger *et al.*, 2007).

1.1.1 Comparaison des profils SSCP normalisés

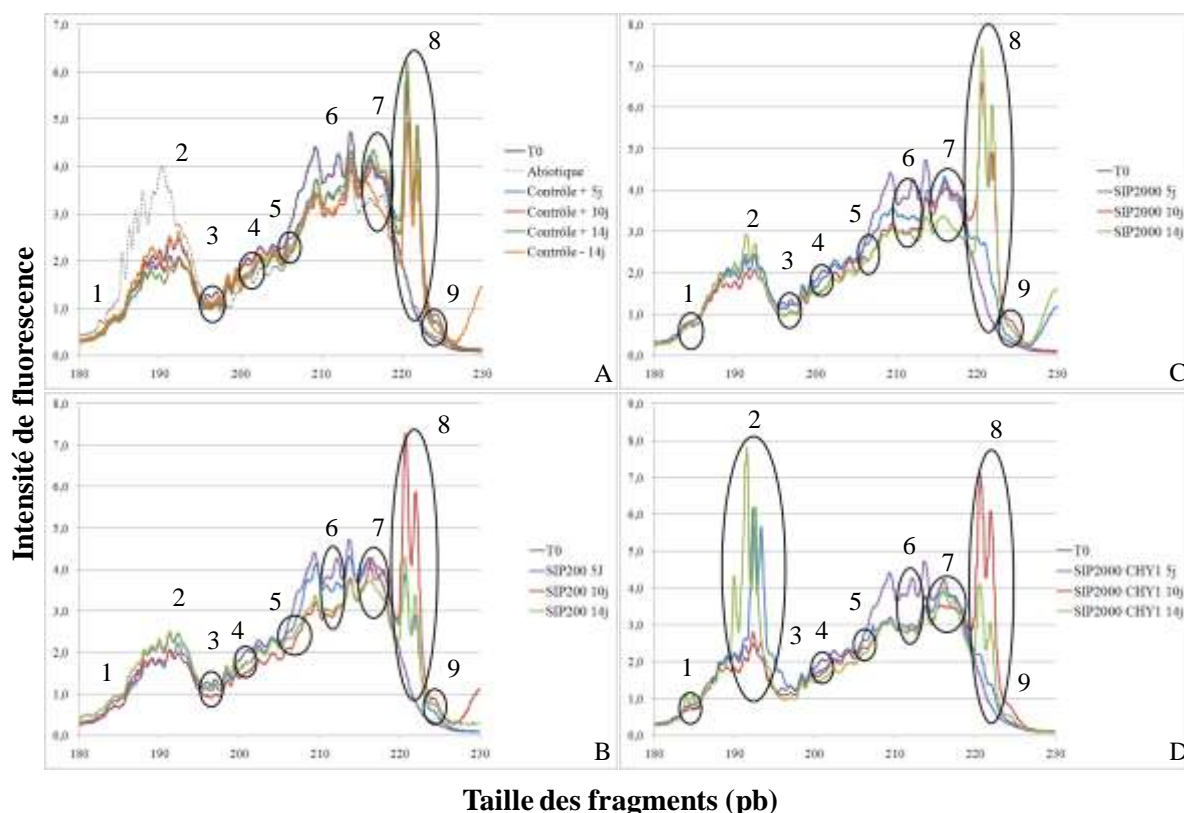


Figure IV.16 : Comparaison des profils de diversité normalisés au cours de l'incubation en microcosmes avec ou sans phénanthrène. Chaque profil de fluorescence est la moyenne de 2 ou 3 réplicats de culture et 2 ou 3 réplicats d'expérience SSCP. Le profil de l'échantillon T0 sert de référence aux quatre groupes d'échantillons ; A : contrôles sans/avec phénanthrène ou abiotique, B : SIP200, C : SIP2000, D : SIP2000 CHY1. Certaines zones où des différences entre profils ont été constatées, sont entourées et numérotées.

Dans un premier temps, les intensités de fluorescence ont été normalisées, et les profils ont été comparés en fonction du temps d'incubation pour chaque type de microcosme. Les résultats ont été rassemblés en 4 groupes (A) les contrôles sans/avec phénanthrène ou abiotique, (B) SIP200, (C) SIP2000, (D) SIP2000 CHY1 (Figure IV.16). Le profil de l'échantillon T0 est commun à tous les groupes et sert de référence correspondant au profil de diversité initial.

A première vue, les groupes A, B, C présentent un profil général similaire qui ne varie pas beaucoup au cours du temps. Les profils du groupe D présentent un pic de forte intensité dans la zone D2 correspondant à l'ARNr 16S de la souche pure CHY-1 (Figure IV.15).

Néanmoins, une lecture plus attentive indique des variations dans neuf zones des profils SSCP (numérotées sur la Figure IV.16). Certains pics sont atténués (zone 3), d'autres disparaissent

(zone 4) ou apparaissent (zone 5) entre le profil initial T0 et les autres profils obtenus à 5, 10 et 14 jours, suggérant une évolution de la diversité bactérienne au cours du temps, indépendamment de l'ajout de phénanthrène.

Des variations de profils ont aussi été observées en fonction de l'ajout de phénanthrène (zones 7, 8, 9), de la dose de phénanthrène ajoutée (zones 1 et 6) ou du temps d'incubation en présence de phénanthrène (zones 2 et 6). Cependant, l'intensité des variations observées ne sont pas quantitativement corrélées aux différences de conditions expérimentales.

1.1.2 Influence du temps et de la dose d'exposition au phénanthrène sur la diversité bactérienne suivie par SSCP

Les profils SSCP ont été comparés deux à deux *via* le logiciel de traitement statistique R afin d'établir des distances relatives entre chacun d'eux. Grâce à une analyse multivariée des données, ces distances ont été ordonnées et représentées dans un espace ordonné à deux dimensions sans corrélation linéaire (Non Metric MultiDimensionnal Scaling ; NMDS) (Ramette, 2007).

Les paramètres « temps d'incubation » et « dose de phénanthrène » ont été confrontés aux variations de profils obtenus, grâce à la fonction « envifit » qui relie les conditions expérimentales aux données de diversité.

Le temps d'incubation est une variable continue en ce sens que même si l'ADN a été analysé à trois temps distincts, l'évolution des communautés bactériennes s'est poursuivie sans interruption durant les 14 jours. Ce paramètre est représenté par un vecteur orienté dans une direction qui reflète son influence sur la diversité bactérienne (Figure IV.17).

La dose de phénanthrène ajoutée est une variable discrète car les trois doses ponctuelles considérées (0 – 33 – 337 ppm) ne constituent pas un gradient au sein d'un même sol. Le facteur est représenté par trois centroïdes ou centres de gravité symbolisant l'influence de la dose sur la diversité (Figure IV.17)

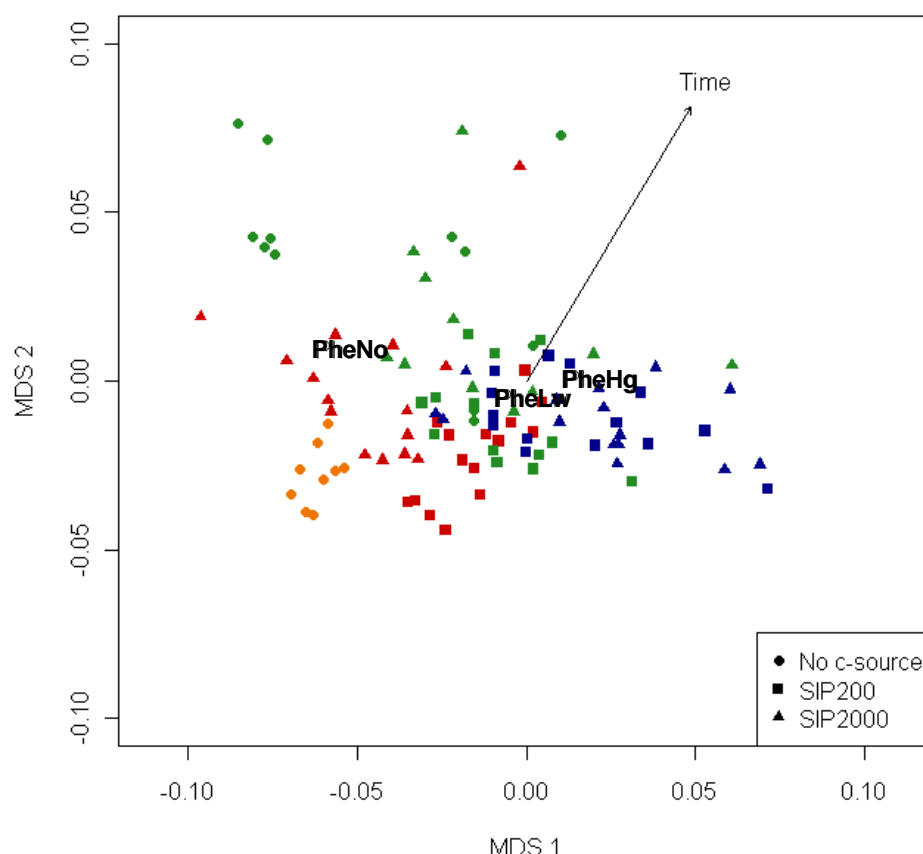


Figure IV.17 : Représentation NMDS des profils SSCP des microcosmes de l'expérience SIP. Les différentes concentrations de phénanthrène ajoutées sont représentées par des symboles : (●) sans phénanthrène, (■) avec phénanthrène (x200), (▲) avec phénanthrène (x2000). Les temps d'échantillonnage sont représentés par des couleurs : T0 orange, 5 j rouge, 10 j bleu, 14 j vert. La variable temporelle est représentée par un vecteur ($p < 0,001$), les taux de phénanthrène ajoutés sont représentés par des centroïdes ($p < 0,002$) : PheNo : pas d'ajout de phénanthrène, PheLw : x 200 et PheHg : x 2000.

Dans les microcosmes exposés au phénanthrène, la variation des profils bactériens est plus marquée entre T0 (orange) et 5 jours (rouge) qu'entre 5 jours et 10 jours (bleu). De plus, aucune variation significative n'est constatée entre 10 jours et 14 jours (vert). Cela suggère que les populations bactériennes sont influencées par le phénanthrène et changent en termes de diversité dans les premiers jours de l'exposition. D'autre part, une variation est observée dans les profils des microcosmes contrôles sans phénanthrène entre 0 et 14 jours (point oranges et verts), suggérant une variation de diversité au cours du temps, indépendante de la présence du phénanthrène.

Les profils de diversité sont plus proches dans les microcosmes traités au phénanthrène (centroïdes PheLw : SIP200 et PheHg : SIP2000) que dans les contrôles non traités (PheNo),

ce qui suggère que le phénanthrène induit des changements dans la structure des populations bactériennes, mais que ces changements sont peu corrélés à la dose de polluant.

L'analyse des profils de CE-SSCP donne une vision globale des changements de la diversité bactérienne d'un sol en fonction du temps ou de certains facteurs environnementaux. Les résultats obtenus ont une valeur statistique indéniable, puisque la rapidité et la reproductibilité des analyses permet de multiplier les mesures et les réplicats. Cependant, l'approche reste essentiellement qualitative et ne donne aucune information sur l'identité des microorganismes responsables des changements détectés. De plus, il arrive souvent que le profil SSCP d'une espèce bactérienne génère plusieurs pics (Ex : *Sphingomonas* sp. CHY-1, Figure IV.15). D'autre part, les profils ont une résolution insuffisante pour rendre compte de la complexité des communautés bactériennes du sol, de sorte qu'un pic est la résultante des contributions de plusieurs ARN 16S, donc plusieurs taxons.

Afin d'identifier les taxons bactériens dont l'abondance varie sous l'effet du phénanthrène dans les expériences SIP, l'abondance des gènes ARNr 16S correspondant aux phylums majoritaires a été évaluée par PCR quantitative en temps réel.

2.2 Effet du phénanthrène sur l'abondance relative des phylums bactériens dominant dans le sol

Les populations bactériennes des sols sont dominées par une dizaine de phylums présents de manière récurrente dans les études de diversité (Janssen, 2006). Ces phylums bactériens peuvent être différenciés sur la base de séquences consensus présentes dans les gènes d'ARNr 16S qui leur sont associés (Fierer *et al.*, 2005; Lopez-Gutierrez *et al.*, 2004; Muhling *et al.*, 2008; Philippot *et al.*, 2010b)

Sept de ces phylums, ainsi que trois classes de Protéobactéries, ont été ciblés par PCR quantitative en temps réel (Figure IV.18). Ces expériences ont été réalisées au laboratoire Microbiologie du Sol et de l'Environnement de l'INRA de Dijon, en collaboration avec Fabrice Martin-Laurent et David Bru.

L'apport de phénanthrène à la dose maximale employée dans cette étude (337 ppm) n'affecte que très peu la population bactérienne totale durant les 14 jours d'incubation en microcosmes

car le nombre de copies de gènes d'ARNr 16S dosés avec les amorces universelles reste stable dans une fourchette de 5.10^5 à $7,74.10^5$ copies par ng d'ADN (Figure IV.18 ; 16S). Toutefois, une diminution significative de l'abondance est constatée après 5 jours.

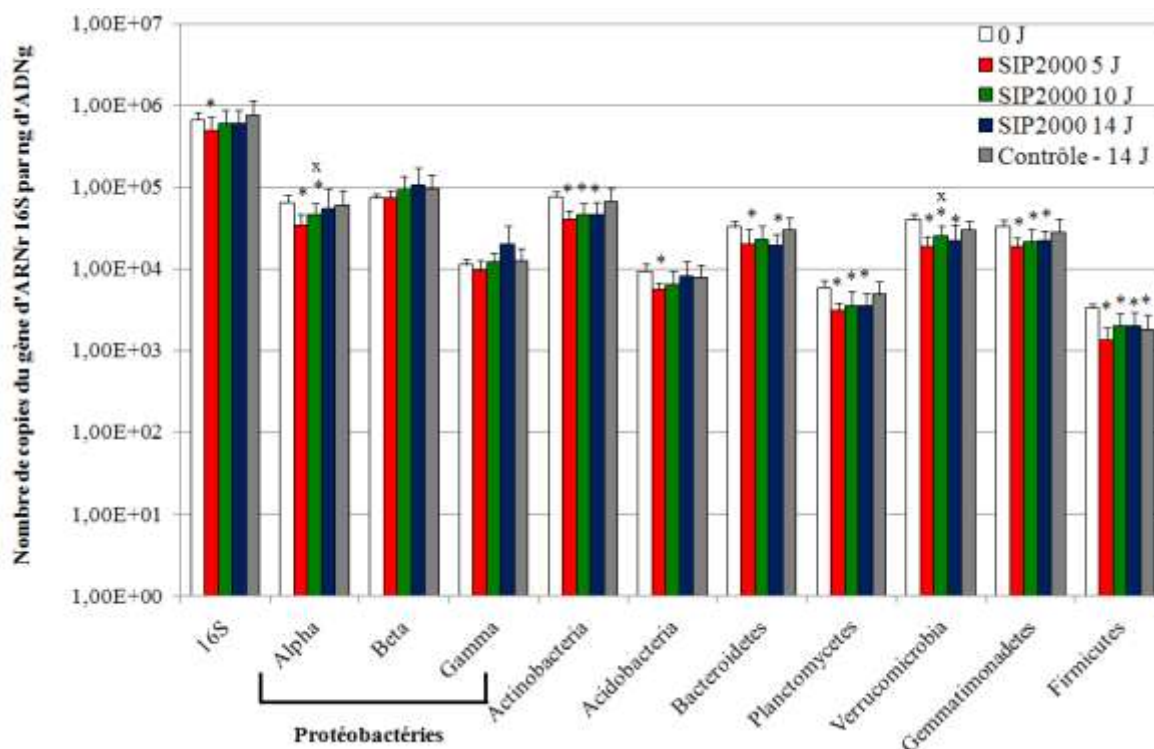


Figure IV.18 : Quantification des gènes d'ARNr 16S totaux et des différents phylums bactériens (ou classes de protéobactéries) ciblés par PCR quantitative en temps réel dans les microcosmes SIP2000 et dans le Contrôle sans phénanthrène à 0 et 14 jours d'incubation. Les valeurs présentées sont des moyennes de copies de gène par nanogramme d'ADN génomique total. Les étoiles représentent les variations significatives par rapport à T0 et les X représentent les variations significatives par rapport à 5 jours, déterminées par le test de Student ($p < 0,05$)

L'abondance des populations d'Actinobactéries, Bactéroidetes, Planctomycètes, Verrucomicrobia, Gemmatimonadetes et Firmicutes diminue légèrement en présence du phénanthrène. Celle des Alphaprotéobactéries diminue dans les 5 premiers jours d'incubation en présence de phénanthrène, mais est compensée au cours des 10 jours suivant.

A l'inverse, l'abondance des Beta- et des Gamma-Protéobactéries varie peu au cours du temps ou augmente légèrement mais les différences ne sont pas significatives. Cependant, si on calcule cette abondance en pourcentage de la population bactérienne totale, c'est à dire par rapport au nombre de copies d'ARNr 16S quantifiées avec les amorces universelles, on constate une augmentation significative de ces deux classes de Protéobactéries dans les 5

premiers jours de l'incubation avec le phénanthrène (Figure IV.19). À 14 jours d'incubation, leur abondance relative n'est toutefois pas significativement différente de celle mesurée dans les microcosmes témoins sans phénanthrène. Cela suggère que des membres de ces deux classes de Protéobactéries ont trouvé des conditions favorables à leur développement pendant les premiers jours de l'incubation dans les microcosmes. Cette période correspond à l'observation du maximum de minéralisation du phénanthrène (Figure IV.13) et à la biodégradation de plus de 90% de l'apport initial (Figure IV.14).

A ce stade de l'étude, on peut soupçonner l'implication des Beta- et des Gamma-Protéobactéries dans la dégradation du phénanthrène. Cette présomption nécessite d'être affinée, notamment par l'étude des populations spécifiquement impliquées dans le métabolisme du phénanthrène, grâce au marquage isotopique (voir le Chapitre III).

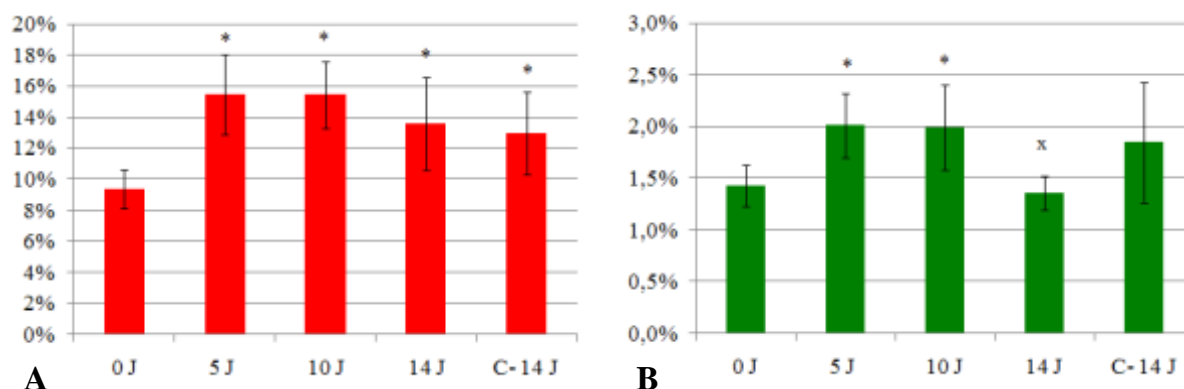


Figure IV.19 : Abondances relatives des Betaprotéobactéries (A) et de Gammaprotéobactéries (B) détectées dans les microcosmes SIP2000 à 5, 10 et 14 jours. Les valeurs sont comparées à celles mesurées dans le sol avant incubation (0 j) et dans les microcosmes après 14 jours d'incubation sans phénanthrène (C-14 j). Les étoiles indiquent des variations significatives par rapport au contrôle 0 j et X représente une variation significative par rapport à 5 jours, selon le test de Student ($p < 0,05$).

Enfin, on remarque que les autres phylums subissent une baisse significative de leur abondance absolue (Figure IV.18) surtout à 5 jours (barres rouges). Cette tendance est significative lorsque l'on regarde les abondances relatives pour les membres de Verrucomicrobia et Planctomycètes (Annexe 4). Cela suggère que le phénanthrène provoque une inhibition sélective de certains taxons bactériens non impliqués dans la dégradation. Excepté pour les membres de Bacteroidetes, l'abondance relative de ces phylums retrouve un niveau proche du contrôle sans phénanthrène après 14 jours, ce qui laisse supposer un retour à l'équilibre des populations bactériennes affectées, suite à la dégradation du phénanthrène.

3. Discussions et conclusions

Dans ce chapitre, nous avons montré que le développement et le métabolisme des populations bactériennes du sol étaient affectés temporairement par l'ajout d'une dose élevée de phénanthrène. L'inhibition de la respiration intervient entre 3 et 7 jours, période qui semble coïncider avec la dégradation de 90% du phénanthrène ajouté. Cependant, une observation plus attentive indique que cette inhibition est décalée de quelques jours par rapport à la période de forte minéralisation du ^{13}C -Phe (entre 2 à 5 jours).

En lien avec cette inhibition métabolique globale, des effets sur la diversité bactérienne ont été observés. En effet, les analyses SSCP indiquent que l'ajout du phénanthrène, quelque soit la dose, influence la diversité surtout dans les 5 premiers jours de l'incubation. D'autre part, on observe une diminution significative du nombre de copies de gènes d'ARNr 16S dans l'ADN total entre 0 et 5 jours. On peut donc conclure que le phénanthrène, à des doses comprises entre 30 et 300 ppm a des effets toxiques sur la communauté bactérienne, qui se manifestent par des changements dans la diversité et une diminution de l'abondance des principaux phylums. Ces changements sont transitoires et suggèrent que la communauté a une bonne capacité de résilience, une fois que le polluant a disparu.

Nos résultats montrent d'autre part que le phénanthrène est rapidement métabolisé dans les 5 jours qui suivent la pollution, sans doute grâce à certaines espèces spécialisées qui se développent entre 2 et 5 jours. De plus, la minéralisation du phénanthrène marqué est d'autant plus rapide que la dose de substrat est élevée. Cela suggère que le métabolisme de ce HAP est d'avantage limité par la disponibilité en substrat plutôt que par la capacité des populations bactériennes à le dégrader.

Par ailleurs, l'abondance des populations bactériennes affiliées aux Beta- et Gamma-Protéobactéries augmente au cours de l'incubation. Etant donnée que leur abondance après 14 jours n'est pas significativement différente en présence (337 ppm) ou en absence de phénanthrène (résultat de qPCR) deux hypothèses sont possibles :

- soit ces bactéries sont capables de se développer plus rapidement que celles des autres phylums sur les substrats naturels du sol

- soit les Beta- et les Gamma- Protéobactéries sont capables d'utiliser le phénanthrène ajouté dans les microcosmes et profitent de cet avantage pour se développer plus rapidement.

L'épisode de contamination au phénanthrène, auquel les microcosmes ont été soumis, met en lumière deux phénomènes opposés et concomitants : un effet toxique et une activité de biodégradation.

Ces deux phénomènes seraient exacerbés par la forte dose de polluant ajouté. Ce substrat est dégradé rapidement par les populations capables de le minéraliser, mais inhibe aussi fortement les populations les plus sensibles, de sorte que les premières se développeraient et que les secondes diminueraient en abondance. Cela pourrait expliquer les perturbations de structures de communautés bactériennes observées par analyse SSCP dans les cinq premiers jours de l'incubation.

Lorsque la majorité du phénanthrène est consommé, les deux phénomènes s'inverseraient. La reprise de la respiration sur les composés ^{12}C correspondrait à une diminution de l'inhibition des populations sensibles, générée soit par des phénomènes de détoxification, soit par une diffusion passive du phénanthrène à l'extérieur des membranes.

Enfin, après 5 jours d'incubation avec la dose la plus élevée de phénanthrène, la minéralisation entre dans une phase lente, traduisant un ralentissement de l'activité des populations actives. La part restante de phénanthrène après 5 jours d'incubation serait moins accessible aux communautés bactériennes laissant supposer deux interprétations possibles :

- soit les populations actives lors de la phase rapide mettent en place des stratégies pour accéder au substrat moins biodisponible, ralentissant leur activité métabolique
- soit d'autres communautés bactériennes spécialistes des substrats adsorbés ou piégés entrent dans le processus de dégradation avec une activité métabolique plus lente.

Les résultats de SSCP et de qPCR sur l'ADN total ne montrent pas d'évolutions significatives des populations bactériennes entre 10 et 14 jours. Des expériences à un niveau d'observation plus fin permettraient de privilégier l'une ou l'autre des interprétations.

Ces hypothèses seront éprouvées dans le chapitre III traitant de la diversité bactérienne spécifiquement impliquée dans la dégradation du ^{13}C -Phe, *via* l'approche de marquage isotopique des populations actives.

Chapitre III. Caractérisation des communautés bactériennes impliquées dans la dégradation du phénanthrène : changements de diversité en fonction du temps et de la concentration de polluant.

Les expériences de minéralisation présentées précédemment ont mis en évidence que le sol d'étude recèle des communautés bactériennes capables de dégrader rapidement le phénanthrène apporté. Nous avons donc émis l'hypothèse que des bactéries spécialisées dans la biodégradation des HAP s'étaient développées dans les premiers jours de l'incubation. Sur la base de l'analyse globale des populations bactériennes des microcosmes, des variations significatives de la biodiversité ont été observées, notamment un enrichissement en bactéries de la classe des Beta- et Gammaprotéobactéries dans les tous premiers jours de l'exposition au phénanthrène. Dans ce chapitre, une analyse détaillée des bactéries actrices de la dégradation de ce HAP a été entreprise grâce au marquage isotopique de ces bactéries *in situ*.

Au cours de cette étude, nous nous sommes fixés trois objectifs : (1) identifier les acteurs majeurs de la dégradation du phénanthrène, (2) déterminer l'influence de la dose de phénanthrène sur la structuration des communautés bactériennes actives, (3) observer le devenir de ces acteurs de la dégradation au cours de l'incubation à mesure que le phénanthrène est dégradé et devient moins biodisponible.

Lors des incubations en microcosmes, les bactéries qui métabolisent le ^{13}C -Phe incorporent l'isotope ^{13}C dans leurs constituants biologiques, en particulier les acides nucléiques.

A partir de l'ADN génomique extrait des microcosmes SIP2000 et SIP200 au chapitre II (10 g), l'ADN marqué a tout d'abord été séparé de l'ADN non marqué sur la base de leur différence de densité. Les gènes d'ARNr 16S, marqueurs de l'identité, ont ensuite été amplifiés, clonés, et séquencés, grâce à une collaboration avec Denis Le Paslier au Génomscope d'Evry. Les acteurs de la biodégradation ont été identifiés par comparaison aux bases de données spécialisées (RDP). Enfin, une analyse statistique des données de séquençage a permis d'évaluer la diversité des communautés bactériennes et de comparer cette diversité en fonction des conditions expérimentales, notamment la concentration en phénanthrène.

1. Isolement de l'ADN génomique marqué ^{13}C par ultracentrifugation isopycnique

La caractérisation des communautés bactériennes impliquées dans la dégradation des HAP, c'est-à-dire celles qui ont métabolisé le ^{13}C -Phe, nécessite de séparer l'ADN métagénomique marqué au ^{13}C de l'ADN non marqué sur la base de leur différence de densité. Cette séparation repose sur une ultracentrifugation sur gradient de chlorure de césium, technique que j'ai introduite au laboratoire, et adaptée à nos conditions expérimentales (paragraphe III.5.3 de Matériel et Méthodes).

1.1 Séparation des formes marquées et non marquées de l'ADN métagénomique

La technique de séparation des formes marquées et non marquées de l'ADN est délicate à mettre en œuvre et sa résolution est l'une des principales limites de l'approche SIP. En effet, la contamination de l'ADN ^{13}C par l'ADN ^{12}C peut conduire à détecter des microorganismes qui ne sont pas impliqués dans la fonction étudiée. Afin de s'assurer que nous parvenions à une séparation satisfaisante dans nos conditions expérimentales, j'ai effectué des expériences préliminaires pour valider notre méthode.

A chaque cycle d'ultracentrifugation, un tube contrôle contenant un mélange d'ADN génomique (^{12}C) de *Escherichia coli* et d'ADN génomique (^{13}C) de *Sphingomonas* sp. CHY-1 cultivé sur ^{13}C -succinate a accompagné les tubes d'expériences SIP.

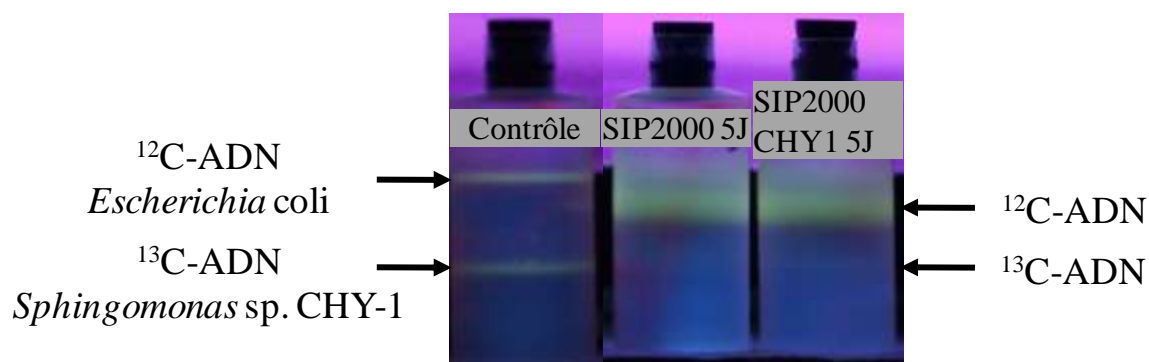


Figure IV.20 : Visualisation sous UV des bandes d'ADN génomique $^{12}\text{C}/^{13}\text{C}$ séparées par ultracentrifugation sur gradient de chlorure de césium. Le complexe d'ADN avec le colorant SYBR Safe est rendu fluorescent par excitation à $\lambda = 254 \text{ nm}$. Le tube contrôle contient un mélange des ADN génomiques de *E.coli* et de CHY-1. Les deux tubes de droite contiennent l'ADN génomique extrait des microcosmes SIP2000 et SIP2000 CHY1 à 5 jours.

A l'issue de l'ultracentrifugation, la révélation sous UV de ce tube contrôle montre deux bandes fines bien résolues correspondant aux deux espèces d'ADN ^{12}C et ^{13}C (Figure IV.20). Dans les tubes contenant l'ADN génomique issu de microcosmes incubés avec ^{13}C Phe, la bande de l'ADN ^{12}C est plus étalée dans le gradient et la bande d'ADN ^{13}C est rarement bien visible. Dans les tubes SIP, les 2 bandes sont plus proches (2 à 3 mm) que dans le tube contrôle (8,8 mm).

Sachant que le gradient est fractionné tous les 150 μl , on s'attend à ce que la concentration maximale d'ADN ^{12}C du tube contrôle soit séparée de la concentration maximale d'ADN ^{13}C par 5 à 6 fractions. Par analogie, cette distance est estimée à 1 à 2 fractions dans SIP2000 5 j et SIP2000 CHY1 5 j.

1.2 Analyse de la résolution de la séparation des deux formes d'ADN

Le gradient a été fractionné et l'ADN génomique de chaque fraction a été purifié par précipitation puis dosé. Le pic d'ADN ^{13}C est séparé du pic d'ADN ^{12}C par environ 5 fractions dans le contrôle (Figure IV.21-A). En revanche, les deux espèces d'ADN des expériences SIP2000 5 j donnent deux pics d'ADN moins bien résolus (1 à 2 fractions) (Figure IV.22-A et B). Ce phénomène résulte de l'hétérogénéité de l'ADN métagénomique dont la composition en GC varie, entre 40 et 70% selon la nature des microorganismes, ce qui augmente la gamme de densité de l'ADN (Lueders *et al.*, 2004).

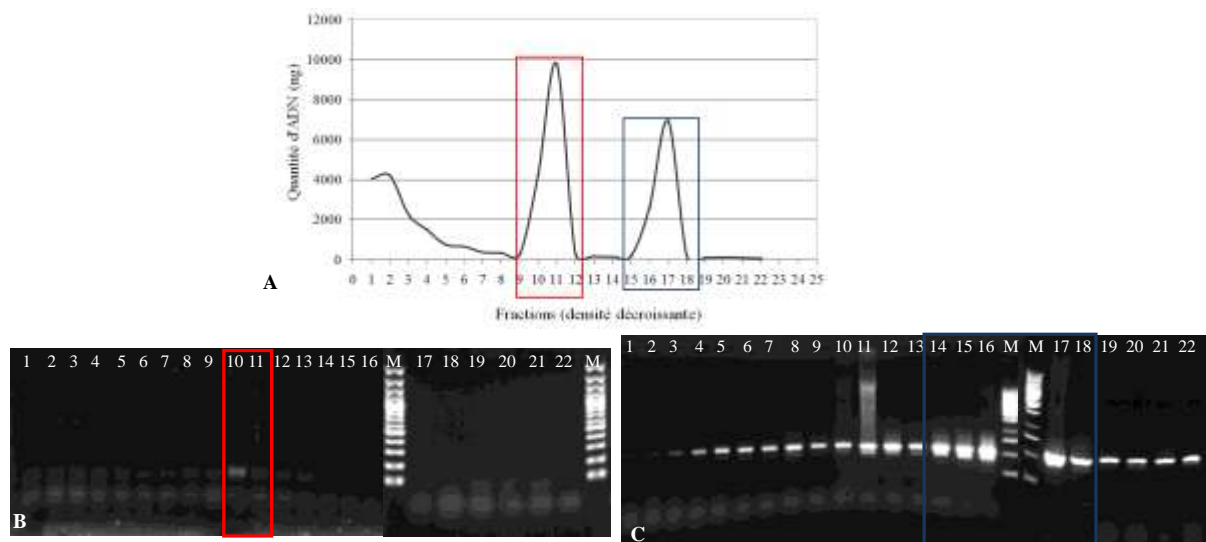


Figure IV.21 : Fractionnement du mélange d'ADN marqué de la souche CHY-1 et non marqué de *E. coli*.
A : dosage de l'ADN génomique dans les fractions purifiées
B : PCR sur les fractions avec les amorces ciblant spécifiquement le gène d'ARNr 16S de *Sphingomonas* sp. CHY-1 (fragment de 85pb)
C : PCR sur les fractions avec les amorces spécifiques de *E.coli* (fragment de 541pb)
 L'encadré rouge désigne les fractions d'ADN ^{13}C , l'encadré bleu les fractions d'ADN ^{12}C .

La résolution des deux pics d'ADN obtenus par fractionnement du gradient a été analysée par PCR. Dans le cas du fractionnement de l'ADN des souches pures, une bande spécifique de l'ARNr 16S de *E.coli* est détectée principalement dans les fractions 14 à 18 (Figure IV.21-C ;) correspondant au pic d'ADN ^{12}C , mais on la trouve aussi dans les autres fractions. Comme attendu, le gène d'ARNr 16S de CHY-1 est spécifiquement détecté dans la fraction 10 correspondant à l'ADN ^{13}C (Figure IV.21-B).

Dans le cas du fractionnement de l'ADN extrait de l'essai SIP2000 CHY-1, on observe que l'ADN de CHY1 est également détecté autour de la fraction 10 (Figure IV.22-B et C ; encadrés rouges). Ceci prouve que la souche CHY-1 inoculée dans le microcosme a bien incorporé le carbone ^{13}C du phénanthrène et que son ADN a bien été séparé de l'ADN non marqué.

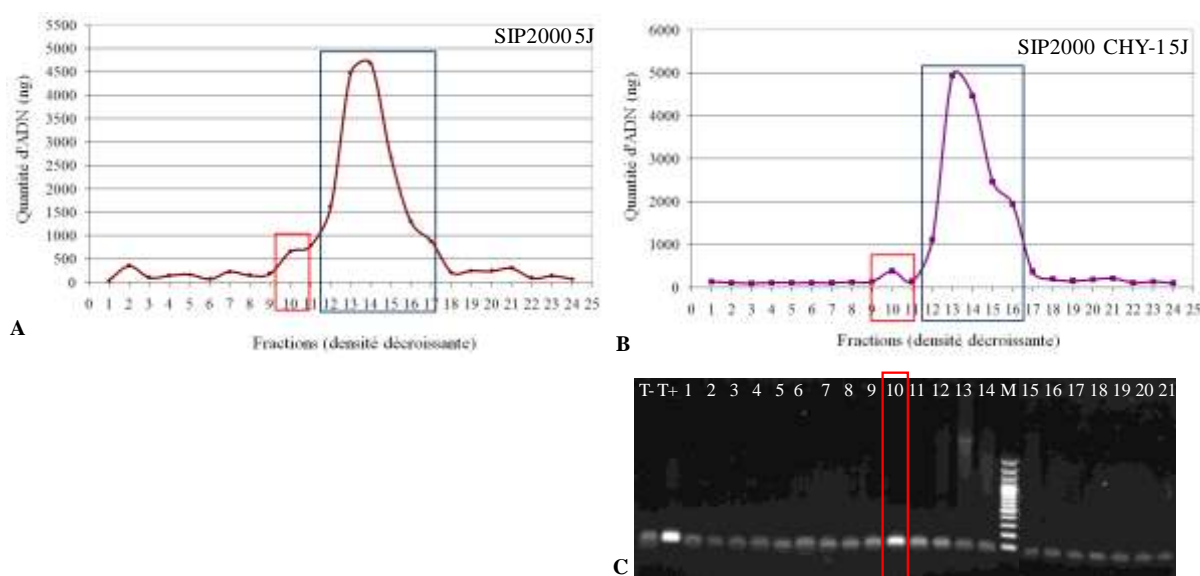


Figure IV.22 : Fractionnement de l'ADN de l'essai SIP2000 à 5 j (A) et SIP2000 CHY1 à 5 j (B).
C : PCR (30 cycles) sur les fractions de SIP 2000 CHY1 5 J avec les amorces spécifiques de *Sphingomonas* sp. CHY-1 (fragment de 85pb)
Puits T- : témoin PCR sans matrice – Puits T+ : contrôle PCR sur l'ADN de *Sphingomonas* sp. CHY-1, Puits 1 à 21 : fractions de l'essai SIP 2000 CHY1 5 J.

Compte tenu de la distribution gaussienne de l'ADN métagénomique, les deux ensembles correspondant aux ADN ^{12}C et ^{13}C se chevauchent. Pour limiter la contamination des ADN l'un par l'autre, les fractions centrales de chaque pic ont été choisies pour constituer les bibliothèques d'ARN 16S relatives aux pools d'ADN marqué et non marqué. Neuf bibliothèques de clones de gènes d'ARNr 16S ont été construites puis séquencées à l'Institut de Génétique

(Génoscope) à Evry (Tableau IV.2). Pour chaque échantillon d'ADN de sol, nous nous sommes appuyés sur une comparaison des librairies de séquences obtenues avec l'ADN ^{13}C et l'ADN ^{12}C pour juger de la spécificité des séquences associées à la dégradation du phénanthrène. Les séquences exclusivement présentes dans les librairies issues de l'ADN ^{13}C ont été considérées comme spécifiques de la fonction ciblée.

Tableau IV.2 : Description des 9 librairies de clones de gènes d'ARNr 16S construites lors de l'expérience de SIP. Les séquences ont été rassemblées en OTUs selon 97% de similarité

Librairies	Nombre de séquences	Nombre d'OTU (97%)
SIP 0 J	231	173
SIP200 5 J ^{13}C	167	132
SIP200 5 J ^{12}C	149	135
SIP200 14 J ^{13}C	353	161
SIP200 14 J ^{12}C	258	233
SIP2000 5 J ^{13}C	424	173
SIP2000 5 J ^{12}C	228	186
SIP2000 14 J ^{13}C	211	135
SIP2000 14 J ^{12}C	28	27
<i>Total</i>	<i>2049</i>	

Les séquences expérimentales ont tout d'abord été comparées à celles de la base de données RDP (Ribosomal Database Project) afin de les affilier à des groupes phylogénétiques de différents niveaux (phylum, classe, ordre, famille, genre). Une analyse statistique a ensuite été réalisée afin de classer les 2049 séquences en OTUs (Unités Taxonomiques Opérationnelles) ou groupes de séquences partageant plus de 97% de similarité. A ce degré de similarité, les OTUs sont considérés comme représentatifs de genres bactériens.

Les librairies ont aussi été comparées entre elles afin d'analyser et de valider statistiquement les résultats obtenus.

2. Structure et dynamique des communautés bactériennes exposées à une forte dose de phénanthrène

Dans un premier temps, nous nous sommes intéressés aux séquences d'ARNr 16S des librairies de SIP2000. A cause de problèmes techniques, un nombre limité de séquences a été obtenu à partir de la librairie SIP2000 14 J ^{12}C (Tableau IV.2). D'un point de vue statistique, ce nombre de séquences a été jugé insuffisant et les données de séquençage ont donc été exclues de l'analyse qui va suivre.

2.1 Identification des principaux acteurs de la dégradation du phénanthrène à 5 jours

La distribution des séquences au sein des phylums bactériens révèle une forte prédominance des Protéobactéries durant les 5 premiers jours d'incubation dans les microcosmes SIP2000 (Figure IV.23-A et B). La classe des Betaprotéobactéries représente notamment 66% des séquences identifiées dans la fraction marquée de SIP2000 5 j. Une analyse à des niveaux taxonomiques inférieurs montre que 4 genres majoritaires (*Acidovorax*, *Rhodoferax*, *Hydrogenophaga*, *Polaromonas*), appartenant à l'ordre des Burkholderiales, dominent les communautés bactériennes impliquées dans la dégradation du ^{13}C -Phe (Figure IV.24-A ; barres rouges). Ces genres bactériens (OTUs 3, 6, 19, 26, 4, 5, 34, 12, 14, 33, 27 ; Figure IV.25) n'ont pas été détectés ou très peu dans les librairies d'ADN non marqué (Figure IV.24 ; barres noires).

Une forte proportion de séquences a d'autre part été affiliée à la famille des Rhodocyclaceae (Rhodocyclales). Ces séquences représentent plus de 30% des Betaprotéobactéries. Bien que les séquences de Rhodocyclaceae soient aussi abondantes dans la librairie SIP2000 5 J ^{12}C , les OTUs obtenus ne sont pas les mêmes dans la librairie d'ADN marqué (OTUs 0, 1, 8 et 17) que dans la librairie d'ADN non marqué (OTUs 28, 51, 71) (Figure IV.25). Cela suggère que certains taxons de cette famille sont impliqués dans la dégradation du phénanthrène. Ces OTUs n'ont pas pu être rapprochés de genres bactériens connus car les séquences auxquelles ils font référence dans les bases de données proviennent de bactéries non étudiées à ce jour.

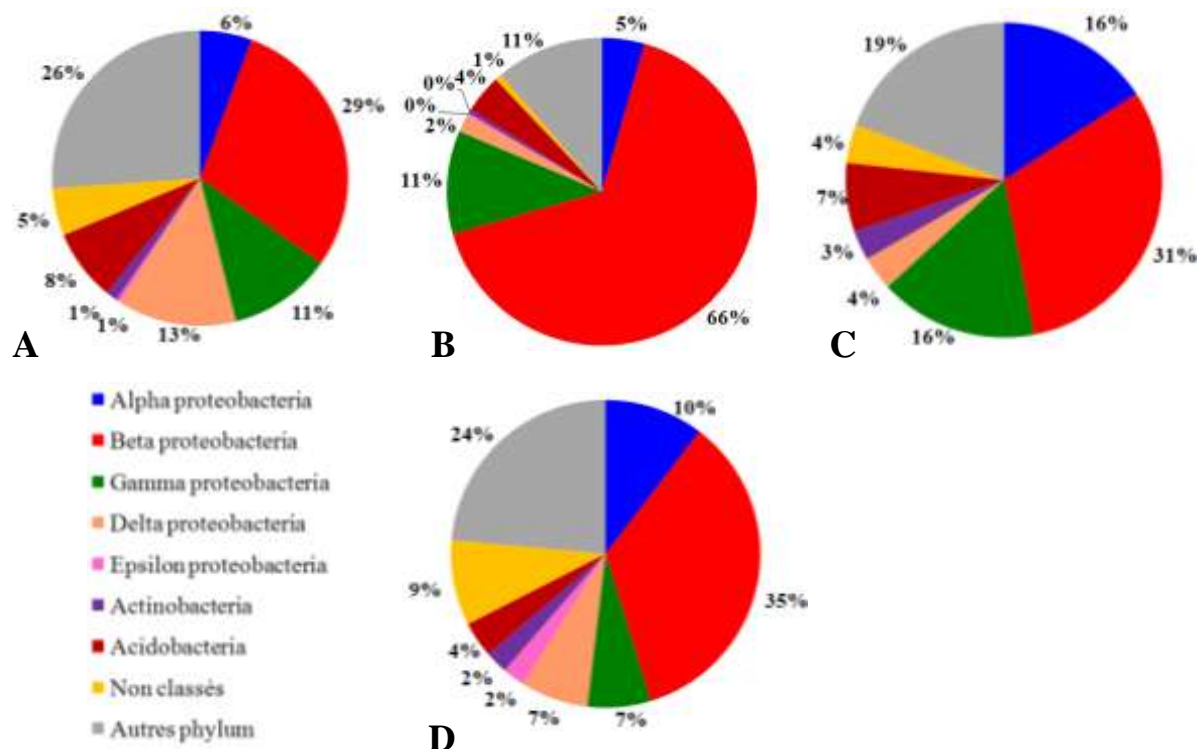


Figure IV.23 : Affiliation des séquences aux phylums les plus abondamment détectés dans les bibliothèques de l'expérience SIP2000. A : SIP 0 J ; B : SIP2000 5 J ¹³C ; C : SIP2000 5 J ¹²C ; D : SIP2000 14 J ¹³C. Les pourcentages représentent le nombre de séquences spécifiques à un phylum rapporté au nombre total de séquences de la bibliothèque. La bibliothèque SIP2000 14 J ¹²C comptant seulement 28 séquences n'a pas été considérée dans ces comparaisons. Sous « Autres phylums » ont été regroupées les séquences affiliées aux taxons suivants : Bacteroidetes, Chloroflexi, Spirochetes, Verrucomicrobia, Cyanobactéries, Gemmatimonadetes, Firmicutes, Planctomycètes, Nitrospira. Sous « Non classés » ont été regroupées les séquences correspondant à des microorganismes non identifiés dans la base.

2.2 Evolution de la diversité bactérienne marquée au cours des 14 jours d'incubation dans les microcosmes SIP2000

A l'issue des 14 jours d'incubation dans les microcosmes SIP2000, les séquences de Protéobactéries restent prédominantes (Figure IV.23-D). De même, les Betaprotéobactéries restent majoritairement détectées. Cependant, un changement significatif dans l'abondance relative des genres dominants est observé par rapport aux résultats à 5 jours.

Au sein des Betaprotéobactéries, la proportion des 5 groupes majoritaires détectés dans SIP2000 5 J ¹³C est bien plus faible (Figure IV.24-A). Cette diminution est corrélée avec la disparition de plus de 90% de phénanthrène constatée dès le cinquième jour d'incubation (Figure IV.14 Chapitre II), suggérant que les Betaprotéobactéries, impliquées dans la

première phase de dégradation du ^{13}C -Phe, voient leur croissance stoppée dès que la biodisponibilité du substrat diminue.

A 14 jours, les Betaprotéobactéries sont dominées par des OTUs affiliés à *Thiobacillus* (Hydrogenophilales ; Figure IV.24-A ; barres bleues), un genre peu représenté chez des acteurs connus de la dégradation des HAP. Peu de séquences affiliées à ce genre bactérien ont été détectée dans les cinq premiers jours de l'incubation.

Comme de nombreuses espèces appartenant au genre *Thiobacillus* sont des autotrophes, on peut se demander si le marquage est lié à la dégradation du phénanthrène ou à l'assimilation de $^{13}\text{CO}_2$ produit dans les microcosmes à partir du substrat marqué. Les séquences les plus représentées (OTU 74 et 123) proviennent majoritairement (66% et 58%) des librairies d'ADN marqué (Figure IV.25), ce qui laisse penser que leur marquage est une conséquence d'une fonction métabolique directement ou indirectement liée à la dégradation du phénanthrène.

Globalement, le nombre de séquences affiliées aux Alphaprotéobactéries augmente légèrement au cours du temps aussi bien dans les librairies d'ADN marqué que non marqué. Toutefois, les Sphingomonadales représentent l'ordre majoritaire dans les librairies marquées (Figure IV.24-B) et un des OTUs (OTU 201) a été exclusivement détecté dans l'ADN marqué.

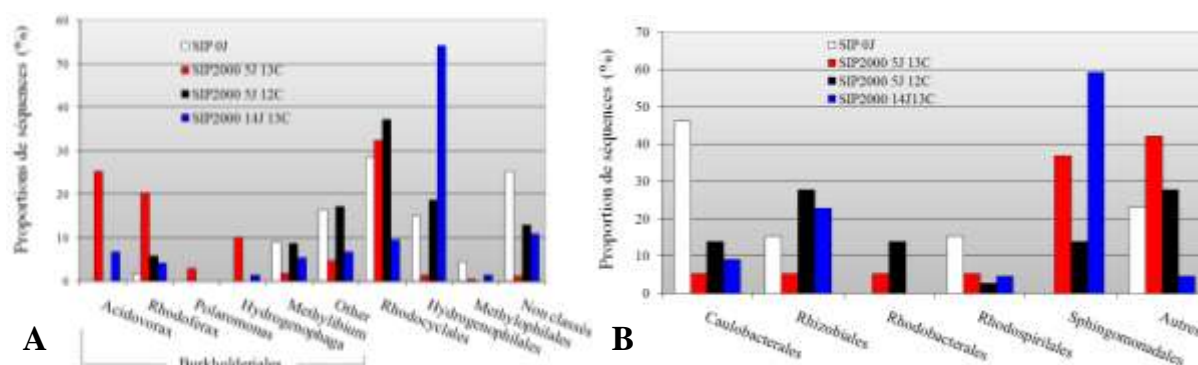


Figure IV.24 : Affiliation des séquences de l'expérience SIP2000 à des niveaux taxonomiques inférieurs pour les Betaprotéobactéries (A) et les Alphaprotéobactéries (B). Les pourcentages représentent les ratios entre le nombre de séquences affiliées à un ordre/genre et le nombre total de séquences de la librairie.

Enfin, dans les librairies d'ADN marqué, 7 séquences sont affiliées à des Actinobactéries, notamment des bactéries du genre *Mycobacterium* (OTU 181) et *Nocardioides*.

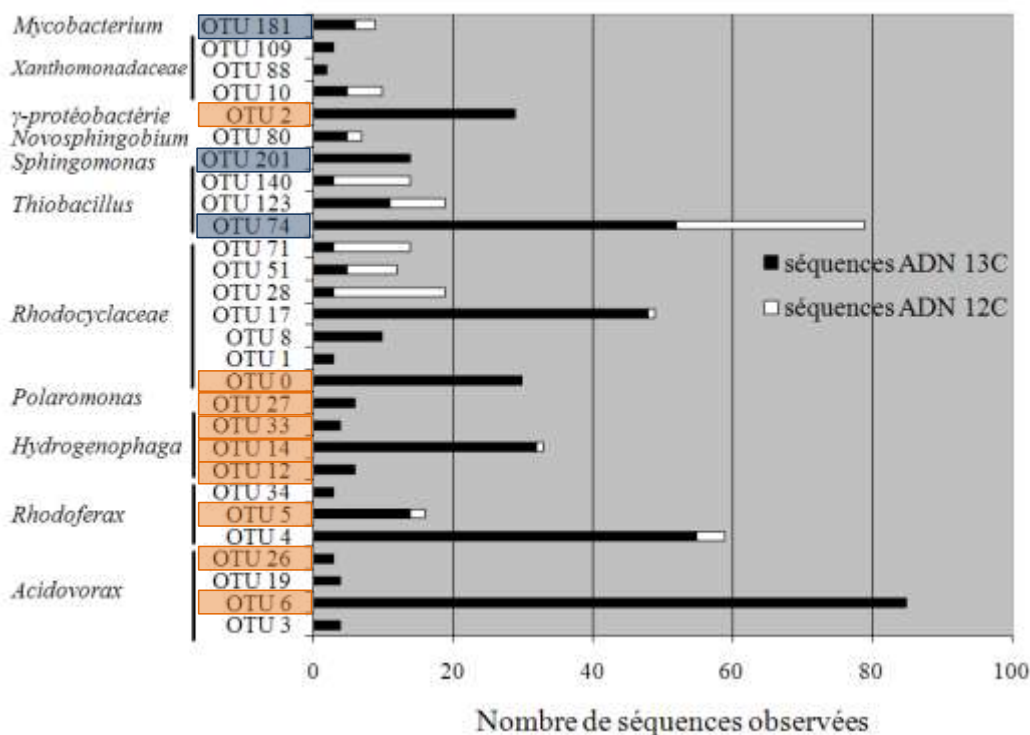


Figure IV.25 : Genres ou familles dominants impliqués dans la dégradation du phénanthrène.

L'analyse a porté sur l'ensemble des séquences des 9 librairies SIP200 et SIP2000. Le diagramme montre une sélection des OTUs les plus représentés par des séquences provenant de librairies ^{13}C (barres noires) ou ^{12}C (barres blanches). Les OTUs de la librairie SIP2000 à 5 jours sont surlignés en orange, et les OTUs majoritaires à 14 jours (SIP200 ou SIP2000) sont surlignés en bleu. Les autres OTUs rassemblent des séquences distribuées dans toutes librairies sans dominance particulière.

De manière générale, les Gammaprotéobactéries sont très peu représentées dans les librairies obtenues. Leurs séquences sont majoritairement affiliées à des bactéries non cultivées et non identifiées. Trois OTUs affiliés à la famille des Xanthomonadaceae (OTUs 10, 88 et 109 ; Figure IV.25) ont été détectés à 5 jours dans les fractions d'ADN marqué et non marqué (OTU 10). L'OTU 2 a été détecté dans l'ADN marqué majoritairement à 5 jours (Figure IV.25) ; cet OTU représente 47% des séquences des Gammaprotéobactéries et 5% des séquences de la librairie. L'identification du genre bactérien auquel appartient cet OTU n'a pas pu être précisé faute de séquences proches dans les bases de données.

L'approche SIP que nous avons mise en œuvre montre (1) que les communautés bactériennes actrices de la dégradation des HAP sont dominées par un petit nombre de taxons essentiellement affiliés aux Betaprotéobactéries, (2) que la structure de ces communautés évolue dans le temps et probablement en fonction de la biodisponibilité du phénanthrène.

2.3 Quantification des phylums bactériens impliqués dans la dégradation du phénanthrène

Les étapes d'amplification, clonage et séquençage comportent des biais qui ont tendance à accentuer la représentation dans les bibliothèques des groupes phylogénétiques les plus abondants au détriment des plus rares (discuté dans le paragraphe I.1.3.2 de l'Etude bibliographique). De ce fait, les données obtenues par séquençage doivent être validées par une autre méthode plus quantitative. C'est pourquoi, une expérience de PCR quantitative en temps réel a été réalisée sur l'ADN marqué. Les phylums ciblés et les modes opératoires sont les mêmes que ceux décrits au chapitre II.

Pour servir de témoins non exposés au phénanthrène, nous avons analysé l'ADN extrait de sol non incubé (T0) et celui d'un microcosme incubé 14 jours sans hydrocarbure ajouté (Contrôle- 14 J). Ces échantillons ont été fractionnés sur gradient et les fractions correspondant à la position de l'ADN ^{13}C ont été utilisées.

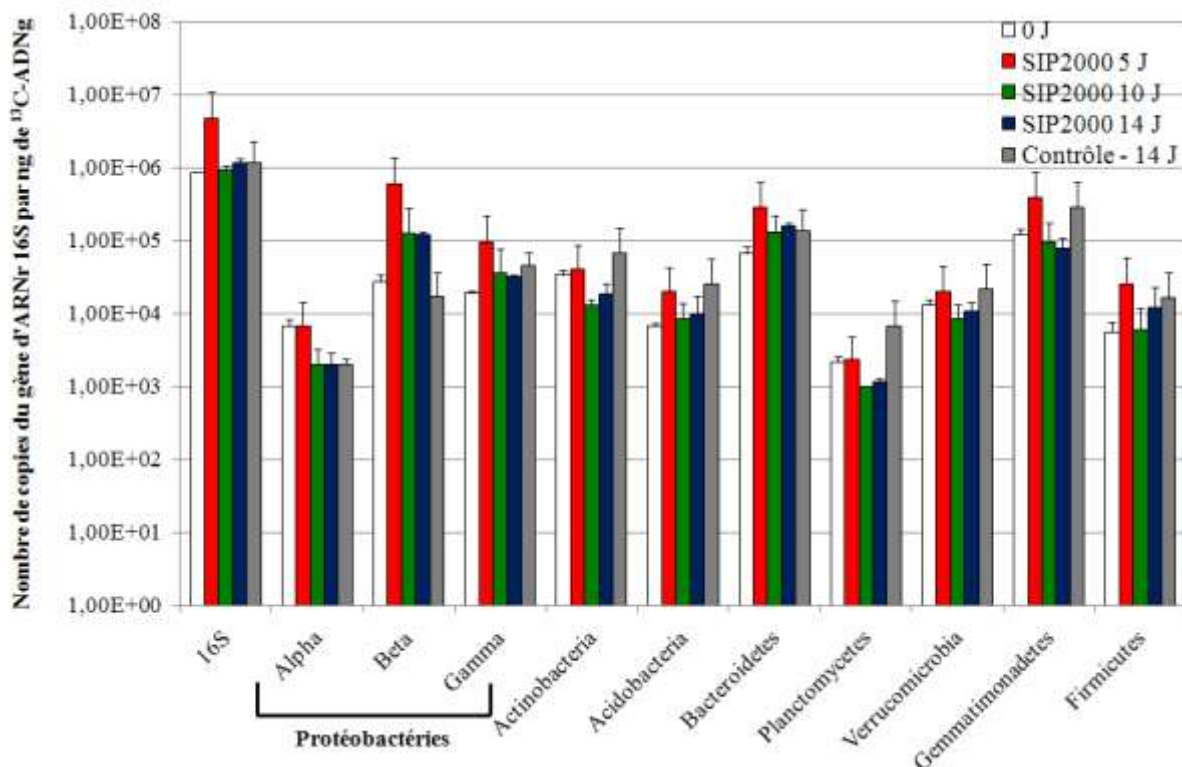


Figure IV.26 : Quantification par qPCR des gènes d'ARNr 16S dans l'ADN marqué ^{13}C des microcosmes SIP2000. Le nombre de copies de gènes d'ARNr 16S totaux (16S) ou spécifiques de phylums bactériens (ou classes de Protéobactéries) a été déterminé à 5, 10 et 14 jours. Des contrôles ont été effectués sur de l'ADN de sol non exposé au phénanthrène extrait au début (0 j) et à la fin de l'incubation (Contrôle- 14 J). Les valeurs sont des moyennes de deux réplicats qPCR sur deux réplicats de culture.

Le dosage des ARNr 16S totaux montre que la population bactérienne marquée a nettement augmenté dans les cinq premiers jours d'incubation en présence du ^{13}C -Phe (Figure IV.26-16S). L'abondance des gènes d'ARNr 16S retrouve un niveau proche du niveau initial (Contrôle 0 j) à partir de 10 jours d'incubation. Cela indique que les bactéries dégradant le phénanthrène se sont multipliées dans les premiers jours de l'incubation.

Cette augmentation de la population marquée résulte essentiellement d'un accroissement significatif de l'abondance des Betaprotéobactéries en présence de ^{13}C -Phe (Figure IV.27-B) et plus particulièrement dans les cinq premiers jours d'exposition (Figure IV.26)

Après 14 jours d'incubation, l'abondance relative des Betaprotéobactéries détectées dans l'ADN marqué reste supérieure à celle du Contrôle sans phénanthrène (Figure IV.27-B).

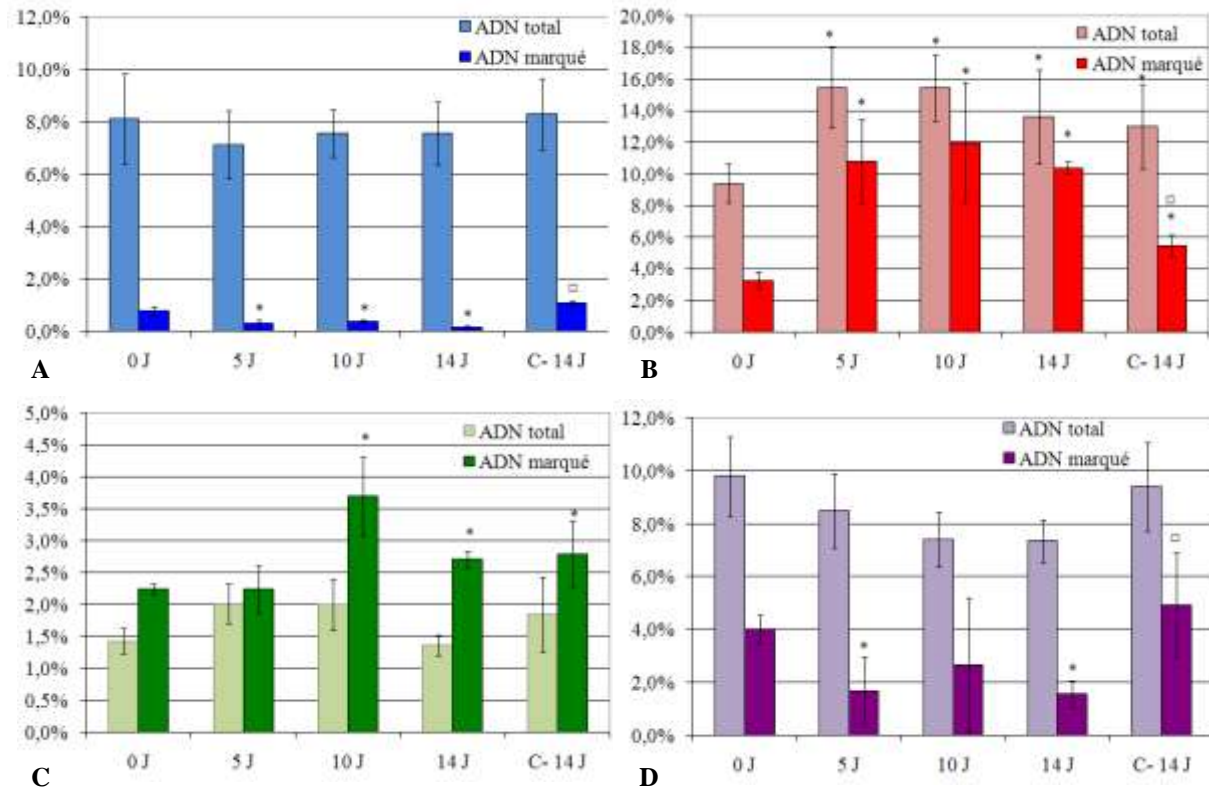


Figure IV.27 : Abondances relatives des Alphaprotéobactéries (A) Betaprotéobactéries (B), Gammaprotéobactéries (C) et Actinobactéries (D) dans les microcosmes SIP2000.

Les abondances rapportées au nombre total de copies d'ARN 16S (amorces universelles) ont été calculées à 5, 10 et 14 jours pour l'ADN marqué (barres foncées) et l'ADN total (barres claires). Deux ADN témoins extraits au temps zéro (0 J) et après 14 jours d'incubation sans phénanthrène (C-14 J) ont servi de références. Les étoiles indiquent une variation significative par rapport à 0 jour et * une variation significative par rapport à 14 jours, déterminées par le test de Student ($p < 0,05$).

Les Betaprotéobactéries sont donc les acteurs majeurs de la biodégradation dans les microcosmes incubés avec 337 ppm de ^{13}C -Phe. Leur abondance relative reste élevée par rapport au témoin sans phénanthrène, même lorsque le substrat ajouté a été consommé à plus de 90%.

L'abondance relative des Gammaprotéobactéries atteint un pic à 10 jours d'incubation. Après 14 jours, elle se maintient à un niveau supérieur au niveau de référence (0 j) mais n'est pas différente de l'abondance mesurée dans le Contrôle- 14 J. Ce résultat indique que des bactéries de cette classe se sont développées sur phénanthrène mais plus tardivement que les Betaprotéobactéries et de manière transitoire. Cependant, ce résultat n'est pas corrélé avec un accroissement du nombre d'OTUs de Gammaprotéobactéries à 14 jours : l'analyse des séquences d'ADN marqué montre en fait que des bactéries de cette classe sont majoritairement détectables à 5 jours (OTU 2, 88 et 109 ; Figure IV.25).

Les Alphaprotéobactéries et les Actinobactéries, connus pour leur implication dans le métabolisme des HAP ne semblent pas tenir un rôle prépondérant dans la dégradation du ^{13}C -Phe à ce niveau d'observation. Le petit nombre de séquences obtenues pour ces deux groupes taxonomiques (Figure IV.24-B) est en accord avec la faible proportion de gènes d'ARNr 16S amplifiés à partir d'ADN marqué (Figure IV.27).

Les autres phylums bactériens ne sont pas spécifiquement enrichis dans les fractions de l'ADN ^{13}C puisque leur abondance relative dans l'ADN recueilli entre 5 et 14 jours est généralement inférieure à celle relevée à T0 (Figure IV.23 et Annexe 4). Néanmoins, l'abondance des Bactéroidetes augmente autour de 10 jours d'incubation. Les séquences relevant de ce phylum dans les bibliothèques SIP2000 à 5 et 14 jours ne sont pas identifiables précisément car elles font référence à des bactéries non cultivées.

3. Structure et dynamique des communautés bactériennes exposées à une faible dose de phénanthrène (SIP200)

La dose de phénanthrène imposée dans les microcosmes SIP2000 était relativement élevée par rapport à la concentration initiale présente dans le sol. Nous nous sommes demandé si une concentration dix fois plus faible, plus proche de celle que reçoit le bassin de rétention de manière chronique, conduit au développement de populations bactériennes similaires à celles trouvées dans les microcosmes SIP2000. La même étude phylogénétique des séquences de gènes d'ARNr 16S a donc été réalisée avec les librairies SIP200

Les Protéobactéries, notamment la classe des Betaprotéobactéries, prédominent également dans les microcosmes SIP200 (Figure IV.28-A). Cependant, après cinq jours d'incubation, on n'observe pas d'enrichissement spécifique de ces phylums dans l'ADN marqué (barres jaunes versus barres grises). En revanche, les Bacteroïdètes et Chloroflexi sont enrichis dans l'ADN marqué à 5 jours (Figure IV.28-B).

Les genres *Acidovorax*, *Rhodoferax*, *Hydrogenophaga*, *Polaromonas*, dominant à 5 jours dans l'expérience SIP2000, sont également bien représentés dans l'ADN marqué des microcosmes SIP200 mais n'apparaissent qu'après 14 jours d'incubation (Figure IV.28-C). Les séquences détectées dans les librairies SIP2000 5 j et SIP200 14 j correspondent aux mêmes OTUs, indiquant que ce sont bien les mêmes bactéries qui prédominent dans les deux cas Figure IV.25).

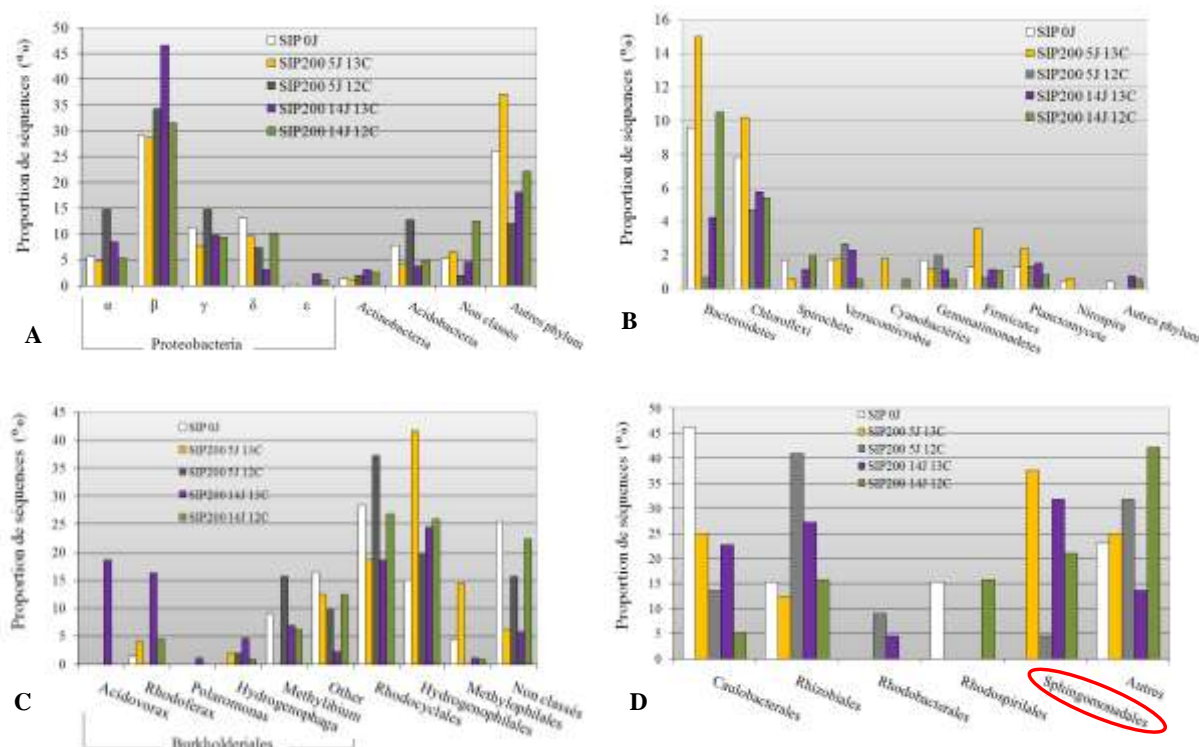


Figure IV.28 : Répartition des séquences de l'expérience SIP200 dans les taxons les plus représentés. A : répartition selon les phylums ; B : détails des phylums notés « Autres phylums » dans l'histogramme A ; C : répartition selon l'ordre et le genre des Betaprotéobactéries ; D : répartition selon l'ordre des Alphaprotéobactéries. Les données sont exprimées en proportion de séquences affiliées à un phylum rapportée au nombre de séquences dans la librairie (A et B) ou en proportion de séquences affiliées à un ordre ou un genre rapportée au nombre de séquences dans la classe analysée (C et D).

Les OTUs affiliés à l'ordre des Rhodocyclales, prépondérants dans la librairie SIP2000 5 J ^{13}C (OTU 0 et OTU 17), sont aussi bien représentés dans SIP200 14 J ^{13}C , mais sont nettement moins abondants (1,1% et 3,4% des séquences contre 6,1% et 8,5% des séquences).

En revanche, les séquences affiliées au genre *Thiobacillus* (Hydrogénophilales) sont nettement plus abondantes dans la librairie SIP200 5 J ^{13}C que dans SIP2000 5 J ^{13}C et SIP200 14 J ^{12}C . On peut supposer que ces bactéries assimilent le phénanthrène mais seulement quand la concentration ne dépasse pas un certain seuil au delà duquel le polluant serait inhibiteur.

Comme dans l'expérience SIP2000, les Alphaprotéobactéries, les Gammaprotéobactéries et les Actinobactéries sont moins représentées et leur proportion change peu au cours de l'incubation (Figure IV.28-A).

Les Alphaprotéobactéries marquées sont dominées par les Sphingomonadales (Figure IV.28-D), bien que peu de séquences soient relatives à cette classe (< 10%) dans les librairies ^{13}C .

Dans les librairies d'ADN marqué, on trouve aussi relativement peu de séquences de Gammaprotéobactéries. Après 14 jours d'incubation, un petit nombre de séquences associées à l'OTU 2 (26% des Gammaprotéobactéries, 2,7% des séquences de la librairie) a été détecté dans l'ADN marqué, ainsi que quelques représentants des OTUs 10, 88 et 109 affiliés à la famille des Xanthomonadaceae (2% des séquences de la librairie pour les trois OTUs). Les autres séquences affiliées à cette classe ne sont pas spécifiques des librairies marquées.

Au sein des Actinobactéries, 2,8% des séquences détectées dans l'ADN marqué à 14 jours ont été affiliées à *Mycobacterium* (OTU 181) et à *Nocardioides*.

Globalement, les communautés bactériennes actives sur les HAP en présence de dix fois moins de ^{13}C -Phe sont semblables à celles observées dans les microcosmes SIP2000, tout au moins pour ce qui est des taxons dominants. Cependant, leur développement est plus tardif et le nombre de séquences obtenues pour chaque taxon dominant est inférieur dans les librairies ^{13}C de SIP200.

4. Analyse statistique de l'ensemble des données de séquençage

L'étude de la diversité bactérienne d'un sol et de l'implication des communautés dans une fonction métabolique ciblée nécessite une détection spécifique des acteurs de la dite fonction. Les approches de marquage isotopique (Ex : SIP) permettent d'isoler l'ADN génomique des populations actives sur un substrat d'intérêt sans pour autant recourir aux méthodes dépendantes de la culture, qui couvrent moins de 5% de la diversité des sols. Cependant, l'une des principales difficultés de la mise en œuvre de la méthode SIP réside dans la capacité à séparer nettement l'ADN ^{13}C de l'ADN non marqué. La contamination de l'ADN ^{13}C par l'ADN de bactéries non marquées a été évaluée par comparaison statistique des librairies de séquences obtenues.

La méconnaissance de l'étendue de la diversité réelle des sols introduit une seconde limite à l'analyse de la diversité bactérienne par séquençage des gènes d'ARNr16S. En effet, comme cela a été discuté dans le chapitre d'introduction, le pourcentage de couverture de la diversité d'un échantillon environnemental par les séquences d'ARNr 16S doit être intégré dans l'interprétation des résultats. Les estimateurs et indices de diversité présentés par la suite nous ont permis d'évaluer jusqu'à quel niveau les données obtenues sont représentatives des communautés bactériennes du sol étudié.

4.1 Evaluation de la part des séquences effectivement attribuable aux bactéries impliquées dans la dégradation du phénanthrène

Lors de l'évaluation expérimentale de la séparation des ADN ^{13}C et ^{12}C (partie I de ce chapitre), nous avons vu que les deux espèces d'ADN se répartissaient selon une distribution gaussienne autour de leur densité moyenne. Il en résulte qu'une part de l'ADN ^{12}C contamine l'ADN ^{13}C présent en quantité nettement inférieure. La comparaison des séquences issues des librairies ^{13}C avec celles issues des librairies ^{12}C provenant du même échantillon montre que des séquences affiliées à la même famille ou au même genre (Rhodocyclales, *Thiobacillus* par exemple) étaient présentes aussi bien dans l'ADN marqué que dans l'ADN non marqué (Figure IV.25). Cependant, cette observation seule ne permet pas de quantifier le chevauchement des librairies de séquences.

A l'aide de la plateforme de logiciels mothur (Schloss *et al.*, 2009) les séquences partageant au minimum 97% de similarité ont été groupées en OTUs (unités taxonomiques opérationnelles), indépendamment de leur appartenance à l'une ou l'autre des librairies. Puis la répartition des 859 OTUs dans les 9 librairies a été analysée.

Les librairies ^{13}C (SIP200 5 J ^{13}C , SIP200 14 J ^{13}C , SIP2000 5 J ^{13}C) partagent moins de 20% de leurs OTUs avec les librairies ^{12}C faites à partir des mêmes échantillons d'ADN (Figure IV.29-B), et entre 12 à 21% d'OTUs avec la librairie de référence SIP 0 J (Figure IV.29-A). Cela suggère que la séparation entre les deux types d'ADN par ultracentrifugation est suffisante pour identifier sans ambiguïté des acteurs majoritaires de la dégradation du ^{13}C -Phe.

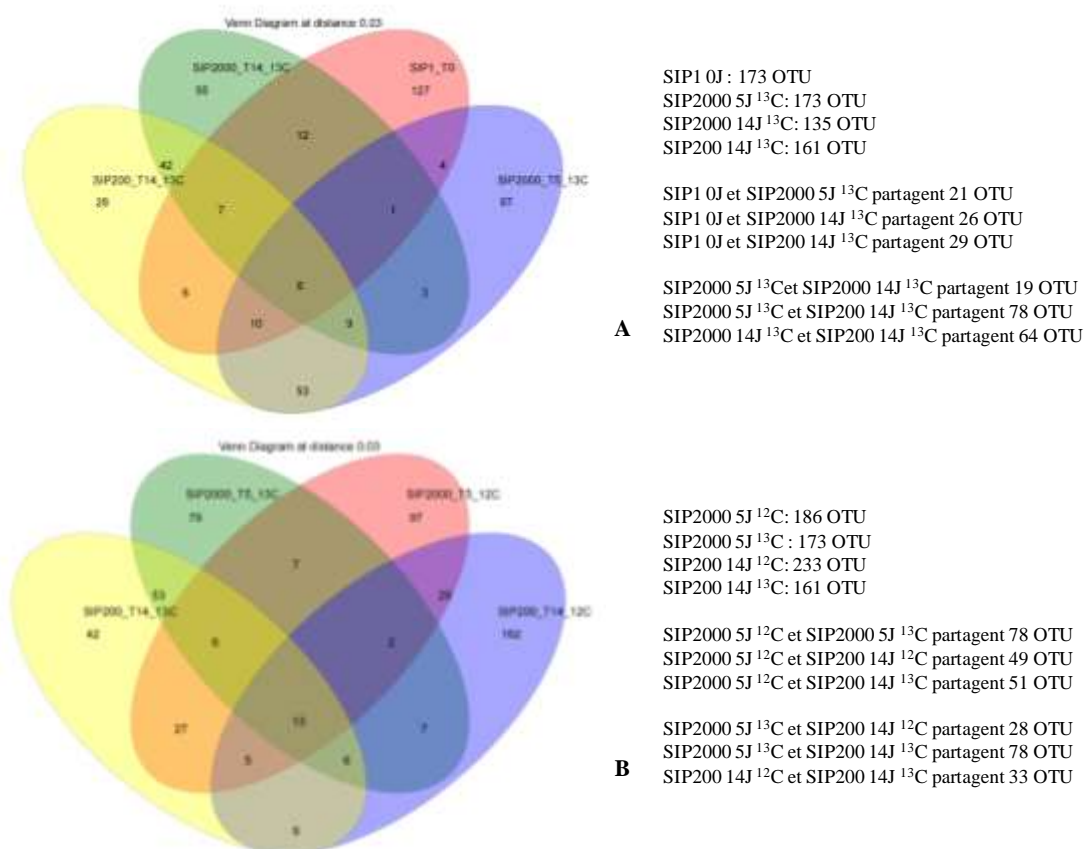


Figure IV.29 : Diagrammes de Venn comparant la distribution des OTUs au sein des librairies de séquences d'ARNr 16S. A : comparaison des librairies issues des fractions d'ADN marqué, SIP200 14 J ¹³C, SIP2000 5 J ¹³C, SIP2000 14 J ¹³C et de T0. B : comparaison des librairies marquées et non marquées se rapportant aux microcosmes ayant la plus forte densité de bactéries dégradant le phénanthrène (SIP2000 5 J et SIP200 14 J). Les aires de recouvrement ne sont pas proportionnelles au nombre d'OTUs partagés. Le pourcentage d'OTUs de la librairie A présents dans la librairie B a été calculé comme étant le nombre d'OTUs partagés par deux librairies (A et B) divisé par le nombre d'OTUs de la librairie A.

Dans les librairies d'ADN marqué des microcosmes SIP2000 et SIP200, moins de 20% des OTUs sont partagés entre T0 et 5 jours et entre 5 jours et 14 jours. Ces résultats montrent que la structure des communautés bactériennes spécifiquement impliquées dans la dégradation du phénanthrène varie en fonction du temps d'incubation.

D'autre part, 33% des OTUs de SIP200 14 J ¹³C se retrouve exclusivement dans SIP2000 5 J ¹³C alors que seulement 5,6% se retrouve exclusivement dans SIP2000 14 J ¹³C (Figure IV.29-A). Ce constat confirme et renforce les résultats présentés précédemment qui montraient qu'après 14 jours d'incubation dans les microcosmes SIP200, la structure des communautés bactériennes marquées est plus proche de celles de SIP2000 à 5 jours qu'à 14 jours (Figure IV.28 ; paragraphe III).

Ainsi, malgré un délai plus long dans le développement des acteurs spécifiques de la dégradation du phénanthrène dans les microcosmes SIP200, les mêmes bactéries dominantes sont observées dans le sol exposé à deux concentrations en polluant différant d'un facteur 10.

4.2 Estimation de la diversité bactérienne des microcosmes par l'analyse des librairies de séquences ARNr 16S

La diversité bactérienne des sols est estimée en moyenne à 10^5 espèces par gramme de sol ((Gans *et al.*, 2005); I. Etude Bibliographique). L'estimation statistique du nombre d'espèces nécessite un échantillonnage représentatif des individus qui la composent. De nombreux outils statistiques ont été développés dans le but de la mesurer et de la prédire. Généralement, la diversité est associée à la notion d'espèce, et on considère que deux bactéries sont de la même espèce quand leurs séquences d'ARNr 16S ont plus de 98% de similarité. Cependant, ce seuil de 98% est arbitraire et peut conduire à une sous-estimation de la diversité, car pour certains genres bactériens comme *Bacillus*, la distinction entre espèces se fait à un seuil supérieur. Volontairement ici, nous éviterons la notion d'espèces et parlerons de taxons ou d'OTUs ayant plus de 97% de similarité de séquence, ce qui correspond à un degré intermédiaire entre le genre (96%) et l'espèce.

Un premier groupe d'outils statistiques, les indices de diversité, décrivent la diversité effectivement observée dans les échantillons. L'indice de diversité de Shannon (H ; Figure IV.30) estime la diversité des taxons contenus dans un échantillon environnemental, en tenant compte du nombre de séquences obtenues pour chaque taxon (Mouillot and Lepretre, 1999). Cet indice a une valeur d'ordre ou d'entropie. Plus sa valeur augmente plus la population observée possède une diversité élevée.

L'indice de Simpson (D ; Figure IV.30), aussi appelé indice de dominance, tient compte du nombre de taxons observés et de la part de chaque taxon dans la population totale. Cet indice donne plus de poids aux taxons les plus représentés. Trois indices de diversité sont déclinés à partir de cette définition. Celui utilisé dans cette étude (1-D) rend compte de la probabilité de sélectionner aléatoirement deux séquences appartenant à deux OTUs différents. Autrement dit, plus cette probabilité est élevée, plus la diversité de la librairie est élevée.

$$H = - \sum_{i=1}^S p_i \log_2 p_i$$

$$D = \sum_{i=1}^S p_i^2$$

Figure IV.30 : Expression mathématique des indices de Shannon (H) et de Simpson (D) ; p_i représente l'abondance relative du taxon i dans un échantillon contenant S taxons.

Les estimateurs de richesse forment un second groupe d'outils statistiques qui tiennent compte de l'effort de séquençage et renseignent sur le nombre de séquences nécessaires pour obtenir une description correcte de la diversité.

Les estimateurs Chao1 et ACE sont des estimateurs de richesse non paramétriques qui tiennent compte soit de la distribution des séquences singletons et doublons (Chao1) soit de la distribution des OTUs abondants (>10 séquences) et rares (<10 séquences) (ACE) (Champreda *et al.*, 2011).

Enfin, la courbe de raréfaction montre la relation entre le nombre de séquences obtenues et le nombre d'OTUs observés. L'extrapolation par une fonction du type Michaelis-Menten (Roesch *et al.*, 2007) permet d'estimer le nombre de séquences nécessaires pour couvrir 95% de la diversité de l'échantillon.

Les outils informatiques de la plateforme mothur sont conçus pour calculer tous ces indicateurs statistiques relatifs aux 9 librairies de séquences obtenues dans cette étude, après regroupement en OTUs à un niveau de similarité de 97%

Les 859 OTUs obtenus sur l'ensemble des séquences sont composés de 534 singletons (62%), 72 doublons et 253 OTUs (30%) possédant au moins 3 séquences. En première analyse, la grande proportion de singletons suggère que la diversité bactérienne des microcosmes est élevée et qu'elle n'a été échantillonnée que partiellement.

L'indice de Shannon proche de 5 (Tableau IV.3) dans toutes les librairies est révélateur d'une forte proportion de taxons. L'indice de diversité de Simpson est quant à lui proche de 1 (Tableau IV.3). La probabilité d'observer deux séquences appartenant à deux taxons différents est grande, démontrant que les taxons rares sont très présents dans les librairies SIP.

Les deux indices de diversité des librairies ^{13}C sont à peine inférieurs à ceux des librairies ^{12}C correspondantes. Cela indique (1) que les acteurs de la dégradation du phénanthrène forment une communauté diversifiée, (2) que ces acteurs sont intégrés dans une population plus large mais qu'ils ne dominent pas la communauté bactérienne du microcosme dans son ensemble.

Tableau IV.3 : Indices et estimateurs de diversité calculés pour chaque librairie au seuil de 97% de similarité de séquences. L'extrapolation des courbes de raréfaction informe sur le nombre de séquences nécessaires pour couvrir 95% de la diversité de l'échantillon.

Librairies*	Nombre séquences	Shannon	Simpson	Chao1	ACE	Extrapolation raréfaction
SIP 0 J	231	4,87	0,99	468	892	9 500
SIP200 5 J ^{13}C	167	4,67	0,95	810	784	13 000
SIP200 5 J ^{12}C	149	4,82	0,99	938	1897	10 500
SIP200 14 J ^{13}C	353	4,32	0,97	748	2144	9 500
SIP200 14 J ^{12}C	258	4,96	0,99	572	956	8 000
SIP2000 5 J ^{13}C	424	3,89	0,95	536	942	7 500
SIP2000 5 J ^{12}C	228	5,01	0,99	739	1981	15 500
SIP2000 14 J ^{13}C	211	4,20	0,96	476	877	7 000

* La librairie SIP2000 14J ^{12}C possédant un trop petit nombre de séquences n'a pas été considérée dans cette analyse statistique.

Les courbes de raréfaction construites avec nos données expérimentales n'atteignent pas une asymptote horizontale (Figure IV.31-A et C), suggérant qu'un plus gros effort de séquençage aurait permis de découvrir des taxons non détectés (au seuil de 97% de similarité). Sur la base d'extrapolations de ces courbes (Figure IV.31-B et C), il aurait fallu déterminer entre 7 000 et plus de 15 000 séquences selon les librairies pour couvrir 95% de la diversité des microcosmes (Tableau IV.3).

Les estimateurs de diversité Chao1 et ACE confirment la richesse taxonomique des échantillons analysés, y compris dans les fractions d'ADN marqué (Tableau IV.3).

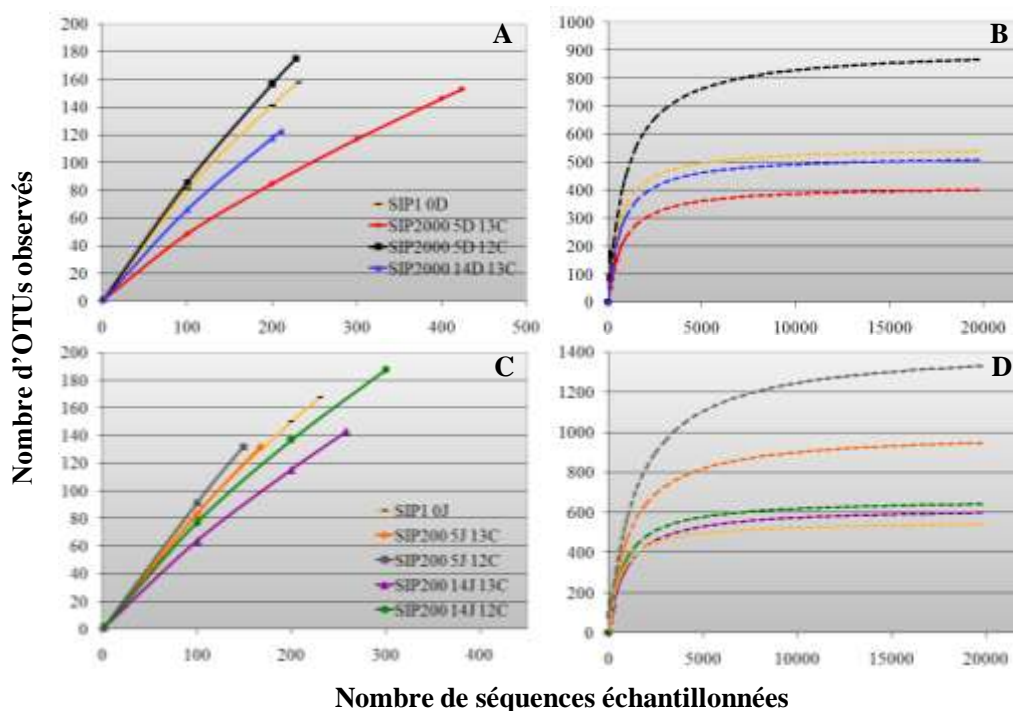


Figure IV.31 : Courbes de raréfactions déduites de l'analyse des séquences d'ARNr 16S issues des librairies SIP2000 (A) et SIP200 (C) et leur extrapolation par la fonction de Michaelis-Menten (B et D respectivement). Pour chaque librairie, les courbes ont été calculées en prenant au hasard des lots de séquences en nombre croissant et en portant le nombre correspondant d'OTUs trouvés pour chaque lot. Le regroupement en OTU a été fait en prenant un seuil de similarité de séquence de 97%.

Nous avons vu (Figure IV.25) que les OTUs les plus représentés provenant de l'ADN marqué correspondaient à des taxons bactériens impliqués dans la dégradation du phénanthrène. En revanche, les OTUs minoritaires (singletons et doublons) également issus de l'ADN marqué ne peuvent pas être associés avec certitude à la dégradation du phénanthrène. Cependant, ces OTUs là contribuent le plus à l'augmentation de la diversité des librairies marquées et surtout à l'estimation du nombre de séquences nécessaires à la caractérisation de la diversité spécifique des acteurs de la dégradation. Ainsi, du fait que notre étude offre un taux de couverture modeste de la diversité, il est probable que nous n'ayons pas détecté des taxons rares susceptibles de jouer un rôle significatif dans la dégradation des HAP.

5. Discussions et conclusions

5.1 Identification des populations bactériennes impliquées dans la dégradation du phénanthrène

L'approche par marquage isotopique mise en œuvre dans ce projet, nous a permis d'identifier les principaux acteurs de la biodégradation du phénanthrène, et ce pour deux concentrations différant d'un facteur 10 (Figure IV.32).

Nous avons mis en évidence que les Betaprotéobactéries jouent un rôle majeur dans ce métabolisme, et parmi elles, les Burkholderiales sont le plus représentées. Nous avons aussi trouvé un taux élevé de bactéries de la famille des Rhodocyclaceae et du genre *Thiobacillus*, alors qu'aucun membre de ces taxons n'a été identifié parmi les bactéries connues pour dégrader les HAP.

Les OTUs affiliés au genre *Acidovorax* (OTU 6, 19, 26) présentent des séquences similaires (> 97%) à des souches identifiées dans un bioréacteur de traitement des HAP, capables de dégrader le phénanthrène (Singleton *et al.*, 2005). Une des souches dominantes de ce bioréacteur, *Acidovorax* sp. NA3, isolée en 2009, est aussi capable de dégrader le naphthalène, le chrysène, le benz[a]anthracène ou le benzo[a]pyrène (Singleton *et al.*, 2009).

Par ailleurs, les autres Burkholderiales identifiées dans notre étude sont apparentées à des espèces de *Rhodoferrax*, *Polaromonas*, *Hydrogenophaga* rencontrées dans des milieux contaminés par les hydrocarbures (Popp *et al.*, 2006) ou sur des filtres de traitement des eaux (Magic-Knezev *et al.*, 2009). D'autres espèces de *Rhodoferrax* (*Rhodoferrax fermentans*) et de *Polaromonas* (*Polaromonas naphthalenivorans*) ont été directement impliquées dans la dégradation du naphthalène lors d'une expérience SIP menée sur un site contaminé (Jeon *et al.*, 2003).

De plus, des membres de la famille des Rhodocyclaceae, représentés par quatre OTUs (OTUs 0, 1, 8, 17) ont été mis en évidence pour leur capacité à dégrader le phénanthrène. Ils sont similaires à deux clones non identifiés, associés, par des approches SIP, à la dégradation

du naphtalène (Figure IV.32, clone 36-9 ; (Bakermans and Madsen, 2002)) ou du pyrène (Figure IV.32, clone Pyr10d2 ; (Singleton *et al.*, 2006)).

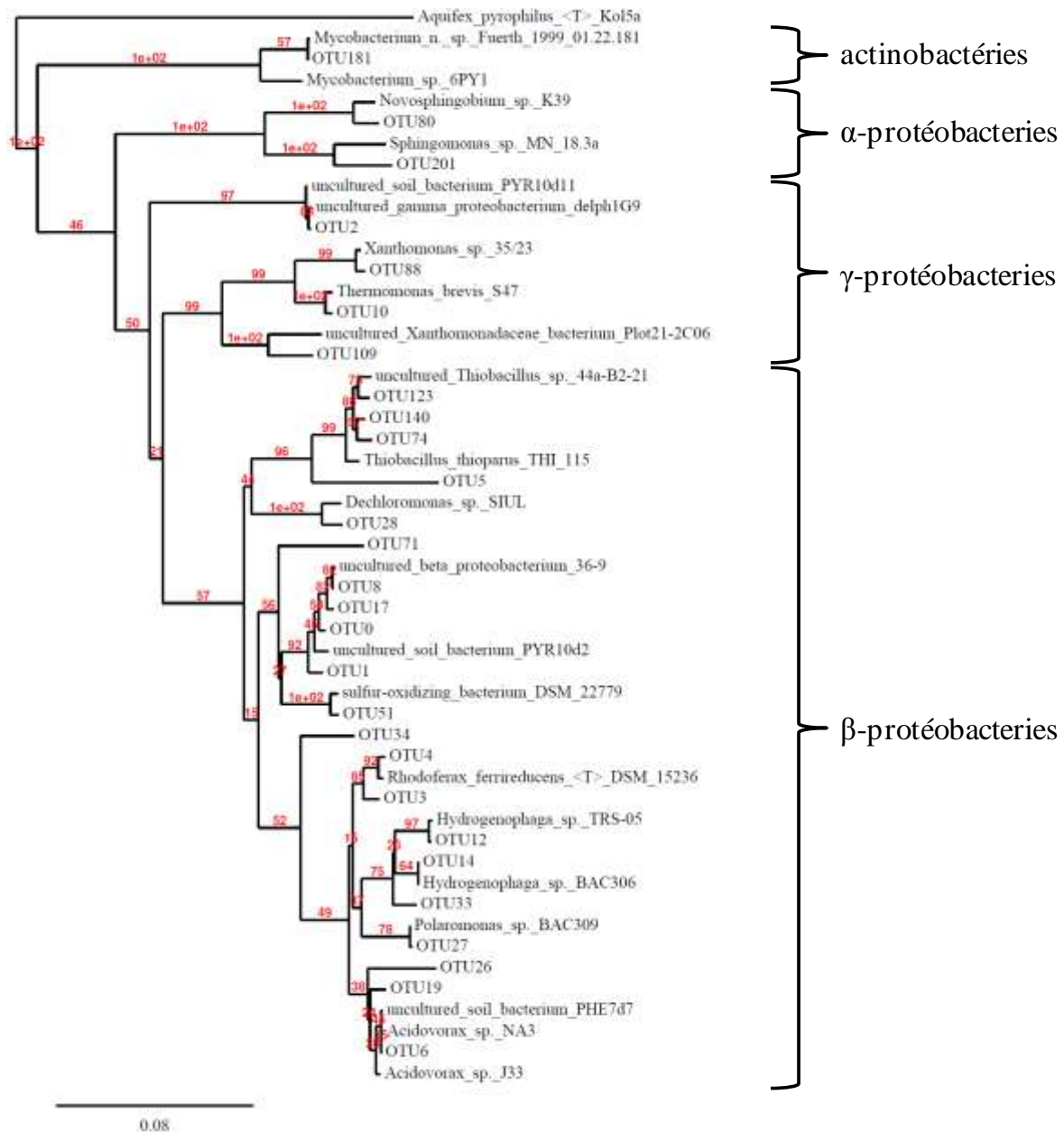


Figure IV.32 Figure 1 : Arbre phylogénétique incluant les OTUs majoritaires associés à la dégradation du phénanthrène et comparaison à des séquences homologues de la base de données RDP.

Novosphingobium sp. K39 (AJ009707) ; *Sphingomonas* sp. MN 18.3a (AJ555475) ; Uncultured soil bacterium clone PYR10d11 (DQ123671) ; Uncultured gamma proteobacterium clone delph1G9 (FM209091) ; *Xanthomonas* sp. 35/23 (AY571841) ; *Thermomonas brevis* S47 (AB355702) ; uncultured Xanthomonadaceae bacterium; Plot21-2C06 (EU193114) ; Uncultured *Thiobacillus* sp. clone 44a-B2-21 (AY082471) ; *Thiobacillus thioparus* THI 115 (HM535225) ; *Dechloromonas* sp. SIUL (AF170356) ; Uncultured beta proteobacterium clone 36-9 (AF351219) ; Uncultured soil bacterium clone PYR10d2 (DQ123667) ; Sulfur-oxidizing bacterium DSM 22779 (AB552842) ; *Hydrogenophaga* sp. TRS-05 (AB166889) ; *Hydrogenophaga* sp. BAC306 (EU130968) ; *Polaromonas* sp. BAC309 (EU130977) ; *Rhodoferrax ferrireducens* sp. DSM 15236 (NC_007908) ; *Acidovorax* sp. NA3 (EU910093) ; Uncultured soil bacterium clone PHE7d7 (AY699590) ; *Acidovorax* sp. J33 (AJ864847) ; *Mycobacterium n. sp. Fuerth 1999* 01.22.181 (AF152558) ; *Mycobacterim* sp. 6PY1 (AM402975) ; *Aquifex pyrophilus* sp.Ko15a (APU17575).

Les autres OTUs de Rhodocyclaceae (OTUs 28, 51 et 71) sont apparentés à *Dechloromonas* ou à des bactéries oxydant le soufre. Dans cette famille, des bactéries dénitrifiantes appartenant aux genres *Azoarcus*, *Aromatoleum*, *Denitratisoma* et *Thauera* ont souvent été décrites pour leur capacité à dégrader des composés phénoliques (Sueoka *et al.*, 2009) ou des alkylbenzènes en conditions anaérobies (Rabus and Widdel, 1995; Rotaru *et al.*, 2010; Suflita *et al.*, 2007). Aucune bactérie capable de dégrader les HAP n'a été décrite dans la famille des Rhodocyclaceae. Si les trois OTUs identifiés correspondent à des bactéries dégradant les HAP, il est peu probable que cette dégradation ait lieu en conditions anaérobies car c'est un processus lent et que la durée de l'expérience SIP n'aurait pas permis un marquage suffisant.

Enfin, notre étude a identifié des OTUs apparentés à *Thiobacillus*, dont les séquences sont retrouvées à plus 50% dans l'ADN marqué. Ce genre bactérien est connu pour sa capacité à oxyder des composés soufrés en autotrophie (Popp *et al.*, 2006). Cependant, il a été démontré que certaines espèces de *Thiobacillus* sont capables d'une activité métabolique hétérotrophe (Kelly *et al.*, 2005). Ce genre bactérien a notamment été trouvé dans un système contaminé par le phénanthrène (Bodour *et al.*, 2003). Nos résultats suggèrent donc que des espèces de *Thiobacillus* sont capables de dégrader les HAP dans le sol.

Les groupes bactériens les plus décrits pour leur capacité à dégrader les HAP, Sphingomonades, Pseudomonades, Actinobactéries (paragraphe I.2.6 de l'Etude bibliographique), sont sous-représentés parmi les taxons que nous avons identifiés *in situ*.

Deux OTUs (OTU 80 et OTU 201) apparentés à des Sphingomonadaceae, sont proches d'espèces décrites, capables de dégrader le chlorophénol (*Novosphingobium* sp. K39 ; (Mannisto *et al.*, 1999)), l'hexane (*Sphingomonas* sp. MN 18.3a) ou le phénanthrène (*Sphingomonas* sp. 3Y). Ces OTUs apparaissent après 14 jours d'incubation quelle que soit la dose de phénanthrène ajoutée. La sous représentation de cette famille est assez surprenante sachant que ces microorganismes sont d'ordinaire très présents dans les milieux contaminés aux HAP (Leys *et al.*, 2004). Elle pourrait s'expliquer par le fait que ces bactéries ont une croissance lente et donc que l'incorporation de ^{13}C n'est détectable qu'après une longue période d'incubation avec le substrat marqué. En accord avec cette hypothèse, l'implication tardive des Alphaprotéobactéries dans la dégradation du phénanthrène a aussi été observée

lors d'études menées sur des sols fortement contaminés par des mélanges d'hydrocarbures (Bordenave *et al.*, 2007; Paisse *et al.*, 2010)

Quatre OTUs de Gammaprotéobactéries (OTUs 2, 10, 88 et 109) ont été détectés dans les bibliothèques SIP2000 5 J ^{13}C et SIP2000 5 J ^{13}C . L'OTU 2, présent majoritairement dans les cinq premiers jours d'incubation avec la concentration de phénanthrène la plus élevée, est proche, par sa séquence, de deux clones identifiés dans la dégradation du pyrène (clone Pyr10d2 ; (Singleton *et al.*, 2006)) ou d'hydrocarbures aliphatiques (clone delph1G9 ; (Peyret *et al.*, 2010)). Les OTUs 10, 88 et 109, majoritairement détectés dans les deux bibliothèques marquées, ont été affiliés à des représentants de Xanthomonadaceae. Cette famille de bactérie est connue pour son implication dans la dégradation des HAP (Hamann *et al.*, 1999; Vinas *et al.*, 2005). Cependant, ces trois OTUs sont très peu représentés dans les fractions d'ADN marqué, suggérant que nos conditions expérimentales n'étaient pas propices à leur développement.

Enfin, le seul OTU affilié à *Mycobacterium* (OTU 181) apparaît en fin d'incubation (14 jours). Il possède 95% de similarité avec le gène d'ARNr 16S de *Mycobacterium* sp. 6PY1, une souche connue pour sa capacité à dégrader le pyrène, le phénanthrène et le fluoranthène (Krivobok *et al.*, 2003). La faible proportion de ce genre bactérien dans nos bibliothèques de séquences pourrait être liée soit à sa croissance lente (marquage tardif), soit à la résistance mécanique de ces bactéries protégées par une paroi très résistante, qui a pu se traduire par un faible rendement d'extraction d'ADN.

En conclusion, nous pouvons dire qu'à la différence des Betaprotéobactéries identifiées comme les principaux acteurs de la dégradation du phénanthrène *in situ*, les Sphingomonadaceae et les Mycobactéries interviennent plus tardivement, quand l'hydrocarbure est en concentration plus faible ou moins biodisponible. Ces bactéries ont une croissance plus lente et seraient spécialisées dans la dégradation des substrats faiblement abondants et peu biodisponibles (Leys *et al.*, 2005).

5.2 Flexibilité de la diversité bactérienne pour répondre à un épisode de contamination

L'apport de phénanthrène dans les microcosmes a stimulé le développement des Betaproteobactéries, mais aussi inhibé le développement des populations sensibles aux HAP, spécialement à la plus forte dose. La diversité bactérienne a donc subi une spécialisation au bénéfice des acteurs de la décontamination du milieu.

Lorsque la quantité de phénanthrène disponible diminue – à 14 jours dans les microcosmes SIP2000 – on pourrait s'attendre à ce que les acteurs prédominants à 5 jours se maintiennent à un niveau élevé, même si leur croissance est ralentie. Or, les résultats montrent que ces populations diminuent nettement dans les neuf jours qui suivent. Cela peut suggérer la présence d'un phénomène de prédation.

Des études récentes montrent que lorsqu'une population bactérienne augmente fortement, cela induit une augmentation du nombre de phages prédateurs associés. Tant que les bactéries sont en développement, l'effet de prédation est marginal, mais lorsque la croissance s'arrête la prédation prend le dessus et la population bactérienne diminue. C'est le principe du Kill the Winner (Rodriguez-Valera *et al.*, 2009). Ce processus naturel de prédation régulerait l'équilibre des communautés bactériennes et préserverait la diversité. Un processus de cette nature pourrait avoir contribué à la diminution des Betaprotéobactéries dégradant le phénanthrène, après consommation de 90% de ce substrat dans les microcosmes.

Le séquençage n'étant cependant pas une méthode quantitative, il est aussi possible que le développement d'autres populations au cours de l'incubation ait un effet de masque sur l'abondance des acteurs de la dégradation à 14 jours, augmentant ainsi l'effet de diminution.

Dans le même temps, on peut penser que les populations inhibées par l'effet toxique du polluant reprennent leur développement lorsque la concentration a diminué. A 14 jours, les résultats de qPCR montrent que l'abondance relative des phylums, hormis les Betaprotéobactéries, tend vers le niveau d'origine (SIP 0 J).

La communauté bactérienne montrerait donc une capacité de résilience naturelle susceptible de préserver la diversité et une faculté d'adaptation à l'exposition chronique aux polluants.

Cette hypothèse est confortée par les analyses statistiques de la diversité. L'indice de Shannon dans les librairies ^{12}C et ^{13}C est plutôt stable et élevé (3,9 à 5). La diversité observée est plus faible dans la librairie SIP2000 5 J ^{13}C ($H = 3,89$), du fait de l'enrichissement en acteurs de la dégradation, mais on observe un retour vers le niveau de diversité d'origine après 14 jours ($H = 4,20$ pour SIP2000 14 J ^{13}C versus $H = 4,87$ pour SIP 0 J).

5.3 Biais liés aux méthodes moléculaires d'analyse des communautés bactériennes

L'exploration de la diversité bactérienne *via* des approches moléculaires suppose la prise en compte des biais inhérents à l'extraction de l'ADN, à l'amplification des gènes d'ARNr16S et aux méthodes de clonage ou de quantification (discuté dans l'étude bibliographique au paragraphe I.1.3).

Des études récentes ont montré que la méthode d'extraction de l'ADN métagénomique influençait la diversité microbienne observée (Delmont *et al.*, 2011; Martin-Laurent *et al.*, 2001). En effet, la nature du sol, la perte de bactéries par adsorption aux particules de sol, ainsi que la lyse incomplète de certains microorganismes affectent le rendement d'extraction des acides nucléiques (Frostegard *et al.*, 1999; van Elsas *et al.*, 2010). Il en ressort que plusieurs cycles ou méthodes d'extraction sont nécessaires pour une analyse exhaustive de la diversité microbienne.

Cependant, une étude récente a montré que si des extractions multiples permettent d'augmenter le rendement d'extraction d'ADN, qualitativement, la diversité bactérienne spécifiquement associée à la dégradation d'un polluant, l'anthracène, n'était pas affectée par le nombre d'extractions et était satisfaisante dès la première (Jones *et al.*, 2011)

La méthode d'extraction d'ADN que nous avons utilisée a donné des rendements moyens de 5 à 20 μg d'ADN/g sol. Si on considère qu'une cellule bactérienne contient 5 fg d'ADN (Leys *et al.*, 2004), ce rendement correspond à l'ADN génomique de 1 à $4 \cdot 10^9$ cellules/g sol. Même en considérant qu'une partie de l'ADN provient d'autres microorganismes, et qu'en réalité le

rendement correspond à un nombre bien inférieur de bactéries, cette valeur est jugée satisfaisante dans la mesure où un sol contient en moyenne 10^9 bactéries/g.

Par ailleurs, les outils moléculaires utilisés pour l'identification et le suivi quantitatif des populations bactériennes influencent les résultats. Ainsi, le choix des amorces d'amplification des gènes d'ARNr 16S et leur caractère plus ou moins universel a une incidence sur l'image de la diversité bactérienne que l'on obtient. De même, les étapes de clonage des gènes d'ARNr16S peuvent accentuer les différences entre taxons abondants et taxons rares. Dans cette étude, des approches moléculaires complémentaires ont été employées pour valider les résultats, en complétant les données de séquençage par des expériences de qPCR.

Des différences ont été constatées dans le cas des Gammaprotéobactéries, peu représentées dans les bibliothèques de séquences, mais détectées en proportion appréciable par qPCR. Toutefois, la principale conclusion de ces travaux, à savoir le rôle majeur joué par les Betaprotéobactéries dans la biodégradation du phénanthrène *in situ* est solidement établie sur la base des analyses de séquences d'ARNr 16S et des mesures d'abondance relative par qPCR.

Chapitre IV. Etude de la diversité des arène-dioxygénases impliquées dans la biodégradation des HAP

Les arène-dioxygénases sont des enzymes qui catalysent l'oxydation initiale de composés aromatiques. Certaines d'entre elles oxydent les HAP. Ces composés étant extrêmement stables, l'attaque initiale est souvent considérée comme l'étape limitante des voies de dégradation bactériennes.

Ce sont des métalloenzymes multicomposantes constituées d'une oxydo-réductase de type flavoprotéine, d'une ferrédoxine (tous deux impliqués dans le transport des électrons) et d'une oxygénase dont le site actif interagit avec le substrat aromatique.

L'oxygénase est une enzyme multimérique associant des sous-unités alpha (50 kDa) et beta (20 kDa) sous la forme d'un complexe $\alpha_3\beta_3$. La sous-unité alpha porte un cluster [2Fe-2S] de type Rieske dans la partie N-terminale, et un centre Fe(II) mononucléaire qui est le site actif de l'enzyme, dans la partie C-terminale. Cette partie constitue le domaine catalytique, responsable de la spécificité du substrat avec notamment une poche hydrophobe dans laquelle le substrat vient se loger (voir le paragraphe I.3.2.3).

Dans ce chapitre, nous avons abordé l'analyse de la diversité des arène-dioxygénases sous l'angle moléculaire et fonctionnel. Dans ce but, nous avons cherché à amplifier, par PCR à partir d'ADN du sol, des séquences correspondant au domaine catalytique des dioxygénases, en ciblant celles qui sont spécifiques de l'oxydation des HAP. Puis, nous avons construit des enzymes hybrides dans lesquelles le domaine catalytique d'une enzyme caractérisée est remplacé par ceux issus d'enzymes des microcosmes, dans le but de doser l'activité enzymatique et de connaître le potentiel catalytique des bactéries du sol.

Pour cette étude, nous avons employé deux enzymes modèles, la dioxygénase de *Sphingomonas* sp. CHY-1 qui a été purifiée et caractérisée au laboratoire, et la dioxygénase de *Acidovorax* sp. NA3, une souche appartenant aux Betaprotéobactéries, et dont les gènes cataboliques ont récemment été identifiés (Singleton *et al.*, 2009).

Deux étudiants de Master II, Laure Malagnoux et Fabien Violet, ont participé à la réalisation des expériences décrites dans ce chapitre.

1. Etude la diversité phylogénétique et fonctionnelle des arène-dioxygénases issues des microcosmes par le biais de constructions d'enzymes hybrides

L'objectif de cette étude était de mettre au point une approche permettant à la fois de cibler spécifiquement les gènes de dioxygénases oxydant les HAP et de sonder leur diversité de séquence, mais aussi d'étudier leur spécificité pour les HAP par la construction d'enzymes hybrides.

L'analyse des gènes d'ARNr 16S des communautés bactériennes dégradant le phénanthrène (Chapitre 3) a montré que les Protéobactéries (Gram⁻) étaient les acteurs majoritaires. L'étude des arène-dioxygénases s'est donc focalisée sur les gènes représentatifs de ce phylum bactérien.

1.1 Principe de la stratégie envisagée

Dans la stratégie envisagée, des séquences codant le domaine catalytique de dioxygénases sont amplifiées à partir de l'ADN marqué des microcosmes à l'aide d'amorces ciblant les gènes appartenant à la famille des HAP-dioxygénases. Ces amorces seront décrites dans le paragraphe suivant (2.1). Les séquences sont ensuite clonées dans un vecteur d'expression contenant les sous-unités complémentaires d'une dioxygénase modèle, en remplacement du domaine catalytique (Figure IV.33).

Ce remplacement doit se faire par insertion des séquences amplifiées dans la phase de lecture du gène de la sous-unité α . Pour faciliter cette construction à partir de séquences variées provenant de l'ADN de sol, nous avons opté pour un clonage PIPE (Polymerase Incomplete Primer Extension ; (Klock *et al.*, 2008); paragraphe III.5.6.5 Matériel et Méthodes) qui permet de cloner des fragments d'ADN obtenus par PCR, en s'affranchissant des étapes de digestion et de ligation. Des gènes hybrides sont ainsi générés.

A ce stade, les amplicons issus de l'ADN de sol peuvent être séquencés afin d'évaluer la diversité phylogénétique des dioxygénases.

Dans un second temps, les dioxygénases hybrides peuvent être co-exprimés avec les gènes des transporteurs d'électrons afin de tester leur potentiel d'oxydation d'une gamme de HAP.

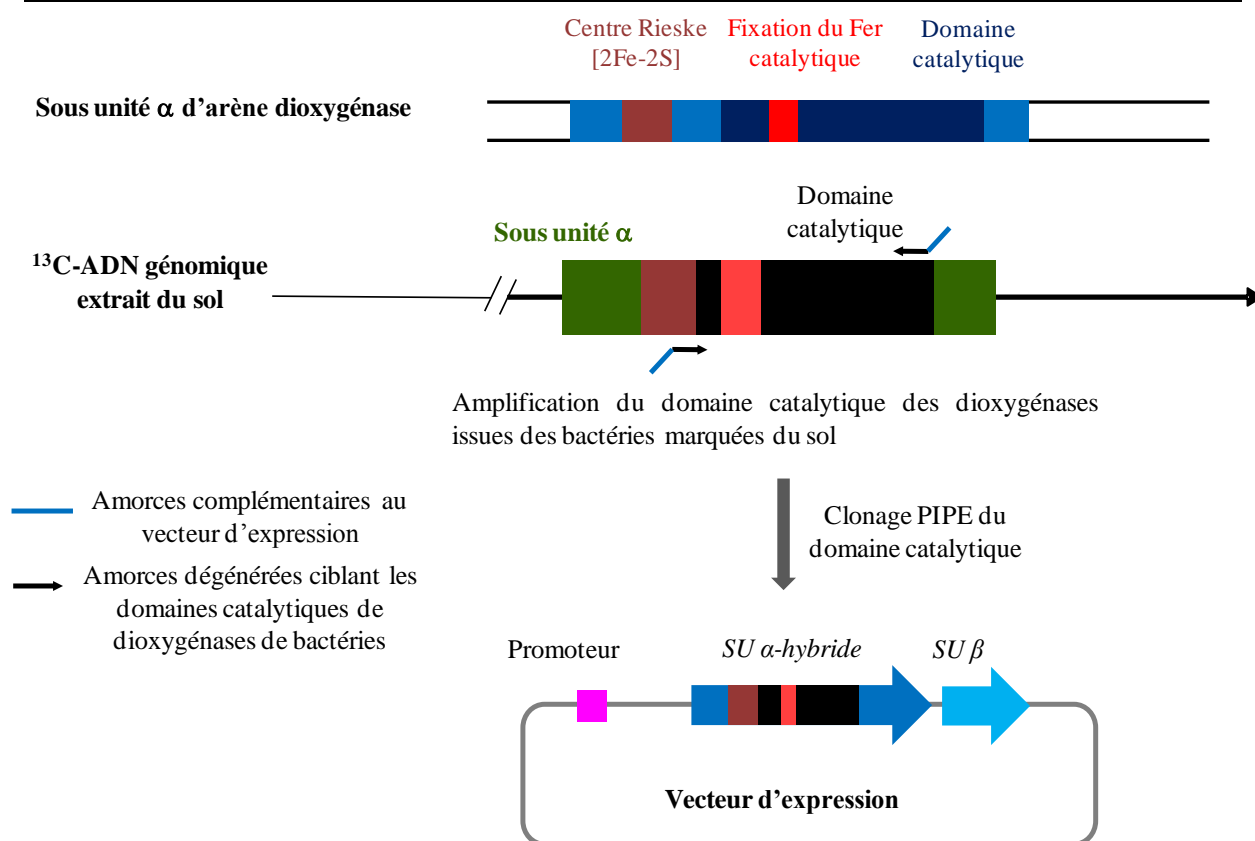


Figure IV.33 Schéma de principe de la construction d'arène dioxygénases hybrides à partir de séquences codant le domaine catalytique d'enzymes du sol. Un schéma plus détaillé du principe du clonage PIPE est présenté dans le chapitre Matériel et Méthodes.

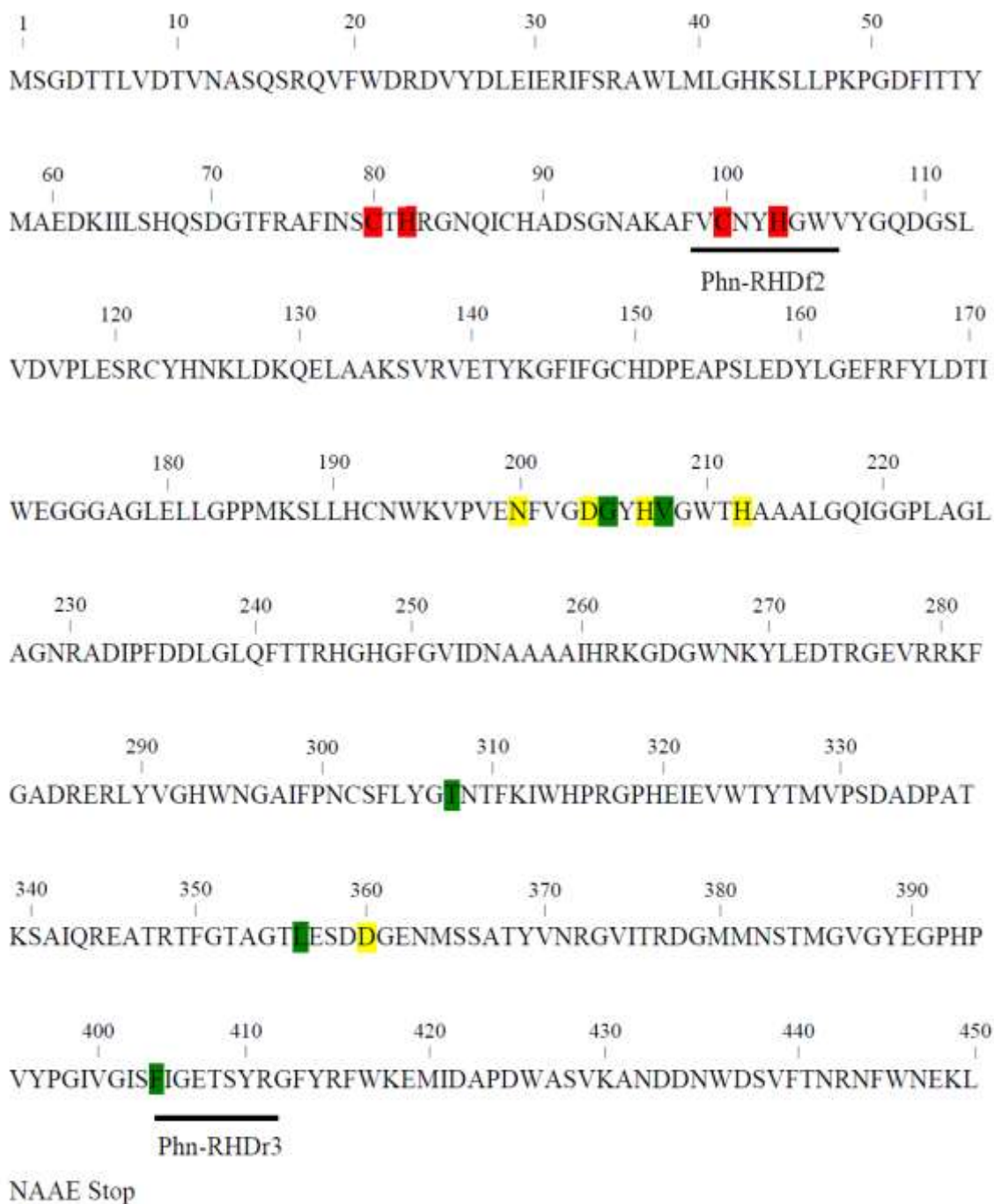
1.2 Difficultés rencontrées et nouvelle stratégie mise en œuvre

La stratégie initiale prévoyait de pouvoir cloner les domaines catalytiques issus de l'ADN des microcosmes à la place du domaine catalytique de PhnI de la souche CHY-1 afin de réaliser à la fois l'analyse phylogénétique et la caractérisation fonctionnelle des enzymes hybrides.

L'analyse de la diversité des dioxygénases suppose que la méthode de clonage employée génère un nombre de clones assez élevé pour obtenir un échantillon de séquences représentatif de la population. Lors de la mise en œuvre du clonage PIPE, nous n'avons pas réussi à obtenir plus d'une soixantaine de clones par cette méthode. Nous avons jugé ce résultat peu satisfaisant.

Nous avons donc choisi de réaliser l'analyse phylogénétique et l'analyse fonctionnelle en deux étapes séparées. Dans un premier temps, les domaines catalytiques ont été clonés dans un vecteur commercial afin d'explorer la diversité des arène-dioxygénases des acteurs de la dégradation du phénanthrène (paragraphe 2 de ce chapitre). Dans un second temps, des gènes

hybrides entre certains domaines catalytiques issus du sol et la dioxygénase PhnI ont été construits afin de doser leur activité enzymatique (paragraphe 3 de ce chapitre).



I : résidus impliqués dans la coordination du centre Rieske [2Fe 2S]

X : résidus impliqués dans la liaison du Fe catalytique

V : résidus impliqués dans la taille de la poche catalytique

— Séquences ayant servi pour définir les amorces d'amplification des domaines catalytiques

Figure IV.34 Séquence polypeptidique de la sous-unité α de la dioxygénase PhnI. Le domaine catalytique amplifié est délimité par les amorces Phn-RHDf2/Phn-RHDr3.

2. Analyse phylogénétique de la diversité des gènes de dioxygénases détectés dans l'ADN marqué.

2.1 Conception des amorces ciblant les gènes de HAP-dioxygénases de type Protéobactéries

Les amorces ont été conçues de manière à cibler le domaine catalytique au sens large, qui s'étend d'un site assez conservé incluant deux ligands du cluster Rieske [2Fe-2S] (nucléotide 291 du gène *phnA1a* de la souche CHY-1) à un site en aval du domaine catalytique (nucléotide 1235 de *phnA1a*) (0).

De plus, nous avons cherché à concevoir des amorces qui soient susceptibles d'amplifier des gènes représentant la diversité la plus large possible, tout en étant spécifiques de HAP-dioxygénases trouvées chez les Protéobactéries. Dans cette optique, les séquences des régions en amont et en aval du domaine catalytique des gènes de dioxygénases de *Sphingomonas* sp. CHY-1 (Demaneche *et al.*, 2004), *Sphingobium yanoikuyae* B1 (Chadhain *et al.*, 2007), *Novosphingobium aromaticivorans* (BphA1f ; (Romine *et al.*, 1999)), *Sphingomonas* sp. LH128 (Schuler *et al.*, 2009), *Cycloclasticus* sp. A5 (Kasai *et al.*, 2003), *Burkholderia* sp. RP007 (Laurie and Lloyd-Jones, 1999), *Acidovorax* sp. NA3 (Singleton *et al.*, 2009), *Alcaligenes faecalis* AFK2 et *Pseudomonas putida* NCIB9816 (Simon *et al.*, 1993) ont été alignées (Figure IV.35). Aucune des deux régions n'étant parfaitement conservée, un consensus a été recherché pour définir les amorces nucléotidiques correspondantes.

Un premier jeu d'amorces dégénérées, tenant compte de toutes les variations de nucléotides, a été testé sur l'ADN génomique de la souche CHY-1 utilisé comme matrice (séquences consensus sur la Figure IV.35). Aucun produit PCR de la taille attendue n'a été obtenu, sans doute parce que chacune des amorces était constituée d'un trop grand nombre de séquences.

Un second jeu d'amorces a été conçu en ne tenant compte que d'une partie seulement des variations de séquence (Phn-RHDf2/Phn-RHDR3). Ce jeu d'amorces, compromis entre diversité et spécificité, a donné satisfaction en permettant l'amplification spécifique du fragment attendu du gène *phnA1a* à partir de l'ADN génomique de CHY-1. Les conditions d'amplification et les séquences de ces amorces sont décrites en Annexe 1.

Amorce amont		Amorce aval	
CHY-1	TTCGTCTGCAATTATCACGGCTGG	CHY-1	TTGGCGAATACATCCTACCGGG
SYB1	TTCGTCTGCAATTACCATGGCTGG	SYB1	TCGGCGAATACCTCTTACCGGG
pNL1	TTCGTCTGCAATTATCATGGTTGG	pNL1	TCGGCGAATACCTCTTATCGAGG
LH128	TTCGTCTGCAATTATCATGGTTGG	LH128	TCGGCGAATACCTCTTATCGAGG
CycloA5	TTTGTATGTAACTATCATGGCTGG	CycloA5	TTGGAGAACTCTTACCGAGG
RP007	TTCGCGTGCAATTATCATGGCTGG	RP007	TCGGGAGAACTCTTATCGCGG
Acido	TTCAATTGCACTTATCACGGCTGG	Acido	TTGGTGAACTCTTACAGAGG
Afaeca	TTCAATTGCACTTATCACGGCTGG	Afaeca	TTGGTGAACTCTTACAGAGG
NdoB	TTTGTCTGCACTTATCACGGCTGG	NdoB	TCGGCGAATACCTCTTATCGTGG
Consensus	TTYRBNTGCAMYTATCAYGGYTGG	Consensus	TYGGHGARACNHBWMHVRNGG
Phn-RHDf2	TTCRYBTGCAAYTATCAYGGYTGG	Reverse	ARCCDCTYTGNDVVWKDBYNCC
		Phn-RHDR3	CCYCKRTAVCWKGTYTCKCCRA

Figure IV.35 : Alignement des régions en amont et en aval du domaine catalytique de dioxygénases connues, pour la conception d'amorces PCR ciblant les gènes d'arène-dioxygénases. Les séquences proviennent de gènes codant les protéines suivantes : CHY-1 : PhnA1a de *Sphingomonas* sp. CHY-1 (CAG17576) ; SYB1: BphA1f de *Sphingobium yanoikuyae* B1 (ABM91740) ; pNL1 ; BphA1f de *Novosphingobium aromaticivorans* (AAD03858) ; LH128 : NahA1f de *Sphingomonas* sp. LH128 (ABW37061) ; CycloA5 : PhnA1 de *Cycloclasticus* sp. A5 (BAC81541) ; RP007 : PhnAc de *Burkholderia* sp. RP007 (AAD09872) ; Acido : PhnAc de *Acidovorax* sp. NA3 (ACG70971) ; Afaeca : PhnAc de *Alcaligenes faecalis* AFK2 (BAA76323) ; NdoB : NdoB de *Pseudomonas putida* NCIB9816 (AAB47591).

2.2 Diversité des arène-dioxygénases de type Protéobactéries

La diversité des gènes de dioxygénases a été étudiée à partir de l'ADN génomique des microcosmes marqué au ^{13}C . Les bactéries dégradant le phénanthrène ont été détectées en plus grande abondance dans le microcosme SIP2000 à 5 jours. L'ADN marqué de cet échantillon a donc été choisi pour ces analyses (Chapitre III.2.1).

Les produits d'amplification obtenus avec les amorces Phn-RHDf2/Phn-RHDR3 ont été clonés dans le vecteur pSTBlue et les inserts de 161 clones ont été analysés par RFLP après digestion par AluI. Un échantillon représentatif des clones ayant un profil différent de celui de *phnA1a* (96 clones) a été sélectionné et analysé par séquençage.

Les séquences nucléotidiques obtenues, dont la taille était conforme à celle attendue (930 à 945 bp), ont été traduites en acides aminés puis comparées aux bases de données NCBI (National Center for Biotechnology Information) par alignement BLAST.

Sur l'ensemble des 96 séquences, 51 formant un groupe appelé dcatdiox A, correspondaient effectivement à des domaines catalytiques de dioxygénases. Ces séquences couvrant plus de la moitié de la chaîne polypeptidique d'une sous-unité alpha (314/450 acides aminés sauf pour

quelques cas où la séquence était plus courte), ont été alignées avec les séquences peptidiques de dioxygénases connues. Un arbre phylogénétique a ensuite été construit à partir de ces alignements pour établir le lien de parenté entre ces séquences, et en déduire la distance relative des dioxygénases issues du sol par rapport à de dioxygénases caractérisées (Figure IV.36). Les séquences du groupe dcatdiox A sont nommées par U3-i ou BBi (i étant un nombre correspondant à la position dans les microplaques de 96 puits).

Les séquences dcatdiox A sont en majorité apparentées à des dioxygénases de la famille des Sphingomonadaceae (Alphaprotéobactéries). Ce résultat suggère que les bactéries de cette famille sont présentes et actives dans les microcosmes SIP2000 bien qu'elles apparaissent comme minoritaires selon l'analyse des gènes d'ARNr 16S (Chapitre III.2).

Les domaines catalytiques détectés se divisent en quatre sous-groupes.

Dans le groupe contenant U3-116, U3-60 et U3-72, 11 séquences sont apparentées à des arène-dioxygénases trouvées chez des espèces connues, comme *Sphingomonas* sp. CHY-1 (PhnA1a_CHY-1), *Sphingobium yanoikuyae* B1 (BphA1f_B1), *Novosphingobium aromaticivorans* F199 (BphA1f_F199), *Sphingomonas* sp.LH128 (NahA1f_LH128), *Novosphingobium* sp. H25 (RHD_ *Novosphingobium*_H25), ou *Agrobacterium* sp. As (BphA_As) (Hiraishi *et al.*, 2002).

Cependant, en majorité, les séquences dcatdiox A sont assez distantes de celles de dioxygénases connues. Ainsi, le sous-groupe contenant la séquence U3-89 (15 séquences), ne présente que 59% de similarité avec la dioxygénase de *Sphingomonas* sp. CHY-1.

De même, dans le sous-groupe contenant U3-16 et U3-124, les 14 séquences partagent au plus 36% de similarité avec la dioxygénase de *Sphingomonas* sp. CHY-1 et 41% avec celle de *Acidovorax* sp. NA3. Les séquences les plus proches sont celles de BphA1b de *Novosphingobium aromaticivorans* F199 (87% d'identité), et AhdA1b de *Sphingobium* sp. P2 (84%). Dans la souche P2, 3 isoformes AhdA1A2 (c, d, et e) ont été décrites comme des salicylate hydroxylases qui interviennent plus tardivement dans la voie de dégradation des HAP (Pinyakong *et al.*, 2003b).

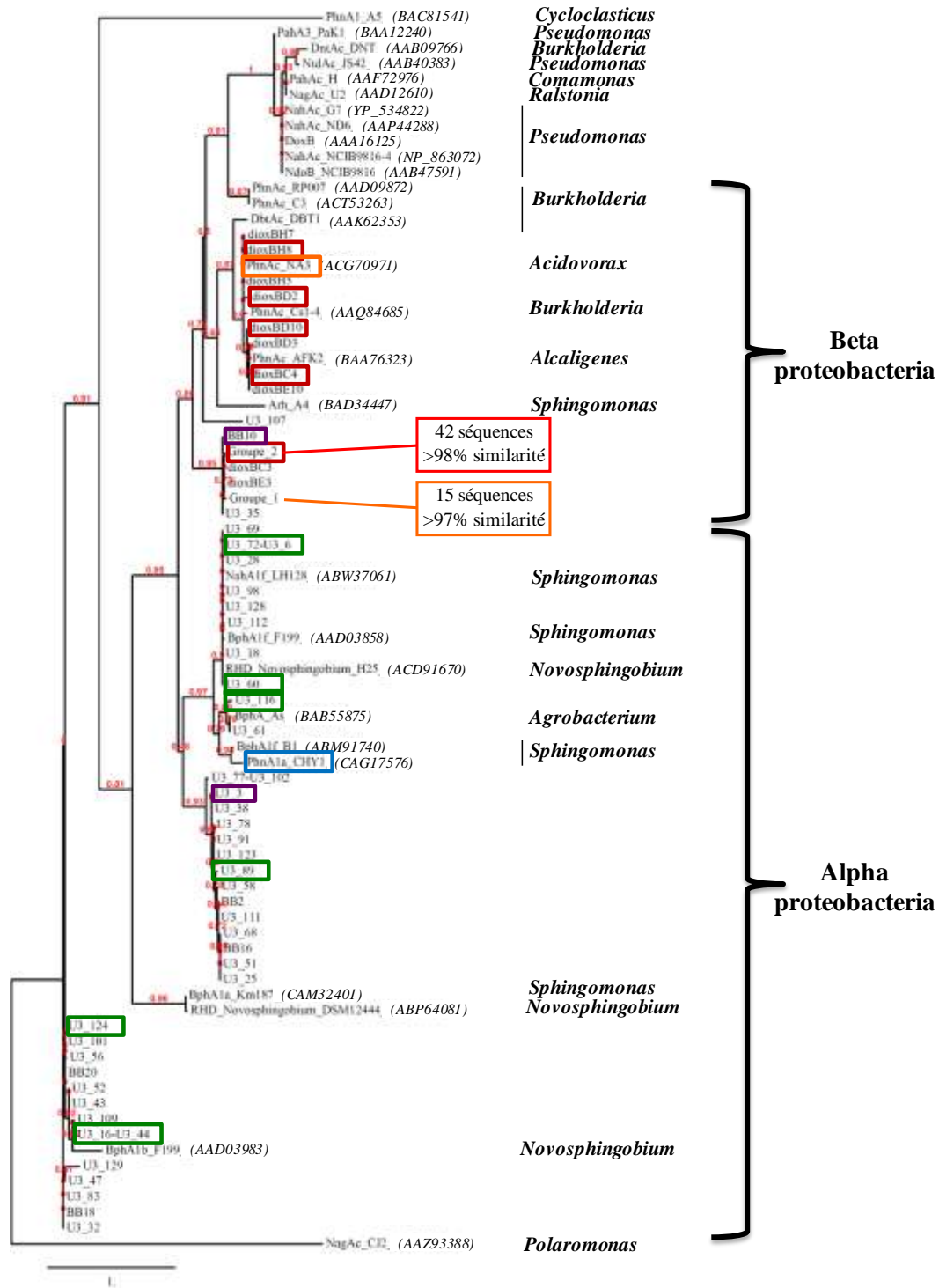


Figure IV.36 : Arbre phylogénétique des séquences de dioxygénases amplifiées à partir de l'ADN génomique marqué extrait du microcosme SIP2000 à 5 J.

Les séquences dcatdiox A sont notées U3-i et BBi, les séquences dcatdiox B sont notées dioxBxi.

Les groupes 1 et 2 correspondent à des séquences présentant plus de 97% de similarité (en acide aminés)

Les encadrés violets, verts et rouges indiquent les séquences qui ont été utilisés pour construire les hybrides.

Les autres séquences sont issues des bases de données. Le nom de la dioxygénase accompagné de son numéro d'accèsion est noté à l'extrémité de la branche correspondante. Le genre bactérien dont elles sont issues est indiqué en italique. La dioxygénase de *Sphingomonas* sp. CHY-1 est encadrée en bleu et celle de *Acidovorax* sp. NA3 en orange.

Quatre séquences, partageant plus de 99% de similarité avec les séquences diox B du groupe 2 (paragraphe 2.3) ainsi que BB10 et U3-35, sont apparentées à des dioxygénases de Burkholderiales.

Grâce au jeu d'amorces Phn-RHDf2/Phn-RHDR3, nous avons donc amplifié des gènes de dioxygénases de séquences variées au sein des Alphaprotéobactéries, la plupart de ces gènes étant distants de ceux répertoriés dans les bases de données.

En revanche, on remarque que les dioxygénases apparentées aux Betaprotéobactéries ne sont représentées que par 4 séquences. Etant donné que cette classe de Protéobactéries est majoritaire parmi les bactéries marquées ^{13}C du microcosme SIP 2000 5 j (Chapitre 3), ce résultat suggère que les amorces favorisaient arbitrairement l'amplification de gènes de Sphingomonadaceae. Nous avons donc entrepris de cibler les dioxygénases de Betaprotéobactéries à l'aide d'un nouveau jeu d'amorces.

2.3 Diversité des arène-dioxygénases de type Betaprotéobactérie

Pour concevoir un jeu d'amorces spécifiques des dioxygénases de Betaprotéobactéries, les séquences de sous-unités α d'oxygénases des bactéries suivantes ont été alignées : *Acidovorax* sp. NA3 (Singleton *et al.*, 2009), *Alcaligenes faecalis*, *Burkholderia* sp. Cs1-4, *Burkholderia* sp. DBT1 (Di Gregorio *et al.*, 2004), *Comamonas testosteroni* (Moser and Stahl, 2001), *Ralstonia* sp. U2 (Zhou *et al.*, 2001), *Burkholderia* sp. RASC (Suen *et al.*, 1996), *Polaromonas naphthalenivorans* CJ2 (Jeon *et al.*, 2006), *Burkholderia* sp. RP007 (Laurie and Lloyd-Jones, 1999), *Burkholderia* sp. C3 (Tittabutr *et al.*, 2011), *Sphingomonas* sp. CHY-1 (Demaneche *et al.*, 2004), et *Novosphingobium aromaticivorans* F199 (Romine *et al.*, 1999). Les amorces ont été réalisées dans les mêmes régions que précédemment (Figure IV.37).

Un premier jeu d'amorces, faites à partir des séquences consensus, a été testé sur l'ADN génomique de *Acidovorax* NA3. Celles ci n'ont pas produit d'amplicon correct sans doute à cause de la trop grande complexité de ces mélanges d'amorces.

Deux jeux d'amorces de moindre complexité ont alors été conçus. Les amorces RHD-Beta-Grp1f et RHD-Beta-Grp1r tiennent davantage compte des séquences de NA3, AFK2, Cs1-4 et

DBT1, alors que les amorces RHD-Beta-Grp2f et RHD-Beta-Grp2r tiennent plus compte des autres séquences.

Seul le jeu d'amorces RHD-Beta-Grp1f et RHD-Beta-Grp1r a donné des résultats satisfaisants en prenant l'ADN génomique de la souche NA3 comme matrice. C'est celui qui a été employé pour rechercher des gènes de dioxygénases dans l'ADN marqué du microcosme SIP2000 5 j.

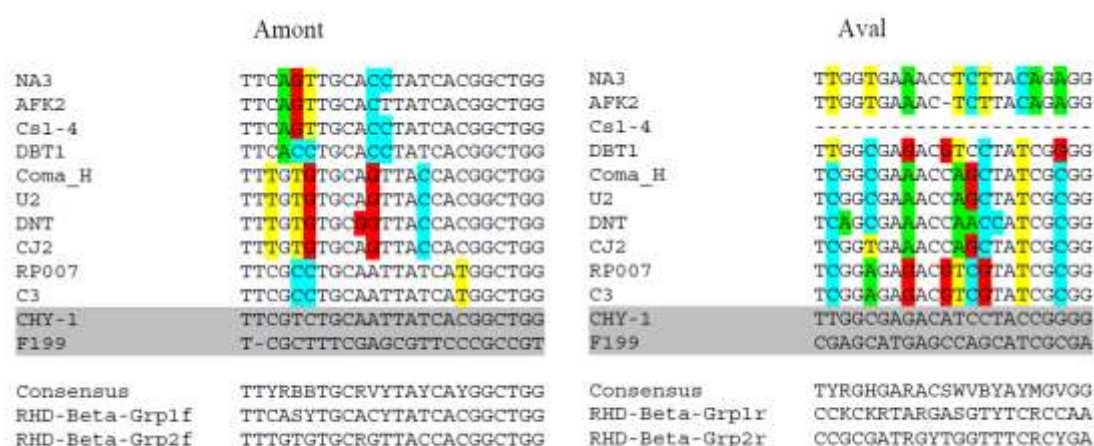


Figure IV.37 : Alignement des régions en amont et en aval du domaine catalytique de gènes de dioxygénases connues de Betaprotéobactéries, pour faire des amorces PCR spécifiques des gènes d'arène-dioxygénase. Les séquences proviennent de gènes codant les protéines suivantes : NA3 : PhnAc de *Acidovorax* sp. NA3 (ACG70971) ; AFK2 : PhnAc de *Alcaligenes faecalis* (BAA76323) ; Cs1-4 : PhnAc de *Burkholderia* sp. Cs1-4 (AAQ84685) ; DBT1 : DbtAc de *Burkholderia* sp. DBT1 (AAK62353) ; Coma_H : PahAc de *Comamonas testosteroni* (AAF72976) ; U2 : NagAc de *Ralstonia* sp. U2 (AAD12610) ; DNT : DntAc de *Burkholderia* sp. RASC (AAB09766) ; CJ2 : NagAc de *Polaromonas naphthalenivorans* CJ2 (AAZ93388) ; RP007 : PhnAc de *Burkholderia* sp. RP007 (AAD09872) ; C3 : PhnAc de *Burkholderia* sp. C3 (ACT53263) ; CHY-1 : PhnA1a de *Sphingomonas* sp. CHY-1 (CAG17576) ; F199 : BphA1f de *Novosphingobium aromaticivorans* F199 (AAD03858).

Les séquences obtenues ont été clonées dans pJET2.1, et 80 clones ayant un insert de la bonne taille (950 pb) ont été analysés par séquençage. Les séquences obtenues ont été traduites en acides aminés puis comparées par BLAST aux bases de données. 65 séquences complètes de domaines catalytiques, formant un groupe appelé dcatdiox B, ont été alignées et comparées au groupe dcatdiox A (Figure IV.36).

Sur la base des similarités observées, l'ensemble dcatdiox B est essentiellement composé de deux groupes de séquences présentant plus de 97% d'identités entre elles : le groupe 1 composé de 15 séquences, et le groupe 2 composé de 38 séquences auxquelles s'ajoutent les 4

séquences de l'ensemble dcatdiox A. Huit séquences très proches de celle de la dioxygénase de *Acidovorax* NA3 ont également été détectées (> 94% de similarité).

Toutes les séquences dcatdiox B sont apparentées à des dioxygénases de Betaprotéobactéries. Cependant, la majorité des séquences ressemblent peu à des dioxygénases connues. Ainsi, les séquences des groupes 1 et 2 ne partagent que 60% de similarité avec la dioxygénase de *Acidovorax* sp. NA3 et au plus 65% de similarité avec l'enzyme la plus proche, une dibenzothiophene dioxygénase de *Burkholderia* sp. DBT1 (Di Gregorio *et al.*, 2004).

En conclusion, les séquences obtenues à partir de l'ADN marqué du sol présentent pour la plupart une parenté avec des dioxygénases connues pour oxyder les HAP. Elles se répartissent en plusieurs groupes phylogénétiques, relativement distants des dioxygénases étudiées à ce jour. Il s'agit donc d'enzymes nouvelles, dont nous avons cherché à connaître le potentiel catalytique d'oxydation des HAP par la construction d'enzymes hybrides.

3. Construction, surexpression et dosage de dioxygénases hybrides

3.1 Construction d'un vecteur d'expression pour la dioxygénase modèle de *Sphingomonas* CHY-1

La naphtalène-dioxygénase de *Sphingomonas* sp. CHY-1 a été choisie comme système modèle pour construire les enzymes hybrides. Elle comprend une composante oxygénase PhnI associée à deux protéines spécifiques, la ferrédoxine PhnA3 et la réductase PhnA4 (Jouanneau *et al.*, 2006). Pour simplifier l'expression d'enzymes hybrides dans *E. coli*, nous avons construit un vecteur portant les 4 gènes de structure de la dioxygénase de CHY-1. Le plasmide pSD8 portant les gènes *phnA1aA2a* sous le contrôle du promoteur T7 a servi de base à cette construction (Figure IV.38). Ce système d'expression est très performant mais produit un taux élevé de dioxygénase insoluble et inactive. Pour palier cet inconvénient, nous avons remplacé le promoteur fort T7 par le promoteur *tac*, afin de se rapprocher d'un système d'expression qui a fait ses preuves (plasmide pSD9, (Demaneche *et al.*, 2004)).

Le promoteur *tac* a été extrait du plasmide pIZpdo2 par coupure BamHI/NdeI, et cloné dans le vecteur pSD8 digérés par BglII/NdeI, remplaçant ainsi le promoteur *T7* (Figure IV.38). Cette construction (plasmide pLMX3) a été vérifiée par séquençage.

Les gènes *phnA3* et *phnA4*, codant les transporteurs d'électrons de PhnI, ont ensuite été amplifiés à partir du vecteur pEB431 avec les amorces phnA43Fbpu/phnA43Rbpu et clonés dans pLMX3 au niveau du site BpuI102I. Une construction ayant les 2 gènes dans la bonne orientation a été sélectionnée par digestion SacI/XbaI (fragment de 1649 pb ; Figure IV.38-C). Le vecteur portant les quatre gènes constitutifs a été nommé pLMX4.

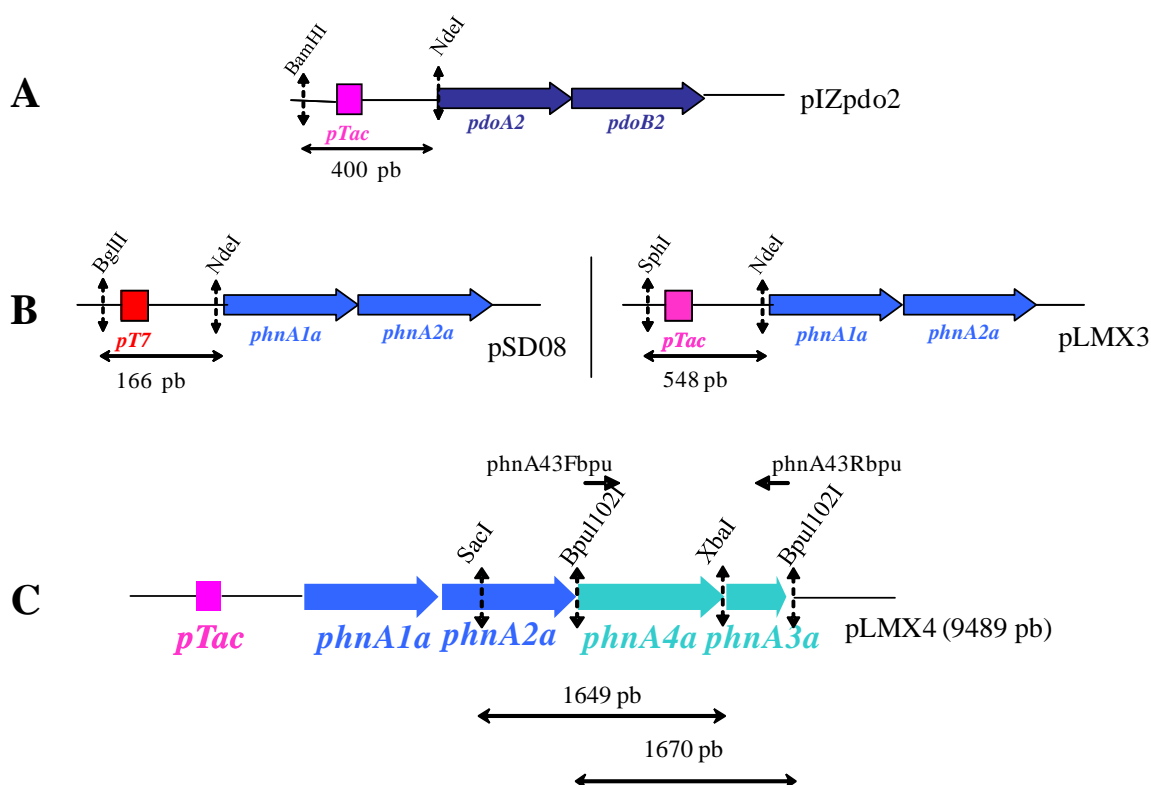


Figure IV.38 : Construction d'un vecteur pour l'expression de la dioxygénase de *Sphingomonas* sp. CHY-1. Les gènes *phnA1a* et *phnA2a* codent pour PhnI, *phnA3* pour la ferrédoxine, et *phnA4* pour la réductase. A : vecteur pIZpdo2 contenant le promoteur *tac*.

B : remplacement du promoteur *T7* par le promoteur *tac*

C : insertion des gènes *phnA4* et *phnA3* dans le plasmide pLMX3 pour permettre l'expression des 4 gènes de la dioxygénase dans *E. coli*.

Les sites de restriction ayant servi aux clonages et à leur vérification sont indiqués par des flèches verticales discontinues.

3.2 Difficultés propres à l'approche expérimentale et mises au point techniques

La démarche expérimentale imaginée pour construire des hybrides est inédite et a donc fait l'objet de nombreuses mises au point techniques. Les principales difficultés auxquelles nous avons été confrontés étaient les suivantes.

Tout d'abord, la technique du clonage PIPE nécessite une amplification PCR non seulement des fragments à cloner (région codant des domaines catalytiques) mais aussi du vecteur receveur. Dans notre cas, le vecteur d'expression que nous avons construit, pLMX4, contenant les quatre gènes *phnA1aA2a*, *phnA3* et *phnA4* (environ 3,8 kb) avait une taille supérieure à 9 kb. L'amplification d'un fragment aussi grand s'est avérée difficile, même en utilisant des ADN polymérases adaptées (KodXL DNA polymérase, Novegen®).

Par la suite, nous avons construit un vecteur plus petit contenant seulement les gènes *phnA1aA2a* (paragraphe 3.3).

Par ailleurs, nous avons constaté que le système d'expression sous contrôle du promoteur *tac* (pLMX3 et pLMX4) n'a pas apporté l'amélioration escomptée : les protéines étaient encore majoritairement produites sous forme de corps d'inclusion. Un autre système d'expression plus adapté a donc été construit comme expliqué ci-dessous

3.3 Construction d'un nouveau vecteur pour la surexpression de dioxygénases hybrides dans *E. coli*

Un vecteur d'expression de taille réduite a été construit sur la base du plasmide pUC18 (2,8 kb) qui présente en outre l'avantage de posséder le promoteur *lac* moins fort que les promoteurs *tac* ou *T7*.

Les gènes des sous-unités α et β de la dioxygénase de *Sphingomonas* sp. CHY-1 (PhnI), portés par le plasmide pSD8, ont été sous-clonés par digestion XbaI-HindII dans pUC18 pour créer le vecteur d'expression pFMA1 (Figure IV.39).

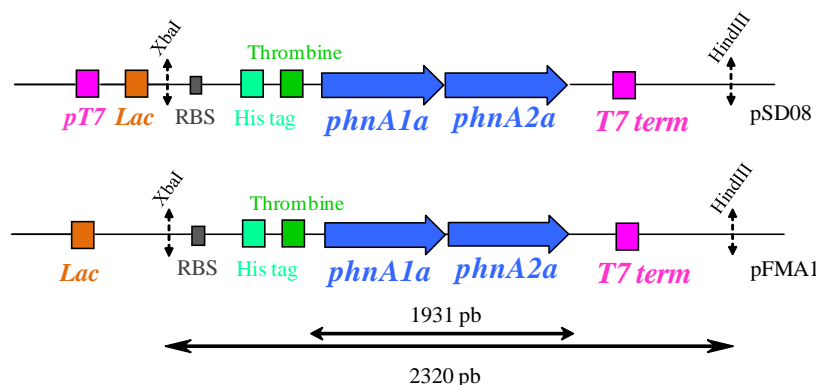


Figure IV.39 : Représentation schématique de la construction de pFMA1, dérivé de pUC18 dans lequel les gènes de l'oxygénase de *Sphingomonas* sp. CHY-1 (PhnI) ont été clonés. RBS : Ribosome Binding Site, His-tag : étiquette poly-histidine, Lac : promoteur *lac*. Les sites de restriction utilisés pour le clonage sont indiqués par des flèches verticales discontinues.

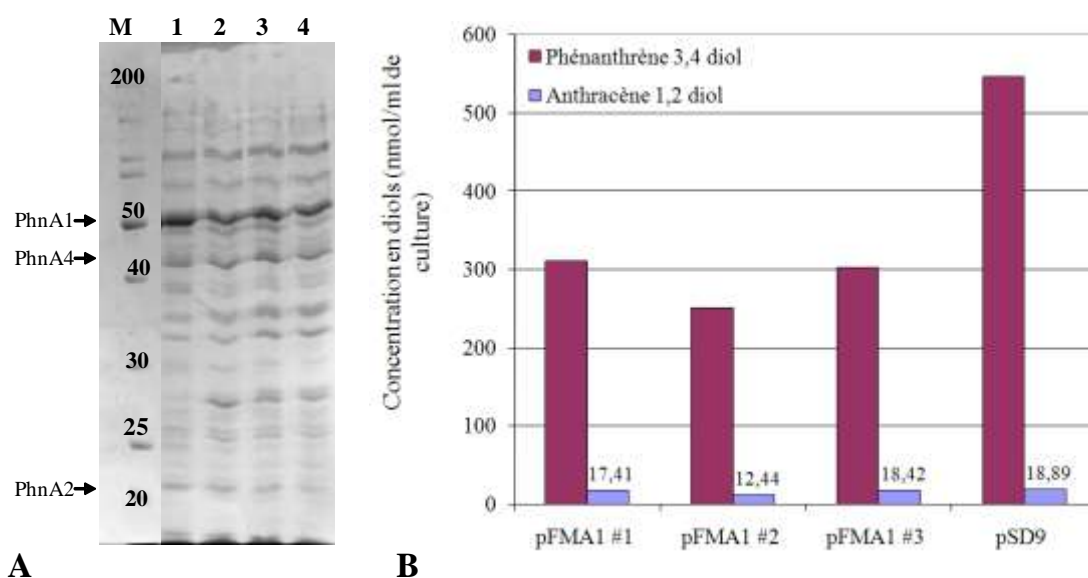


Figure IV.40 : Expression et dosage d'activité de la dioxygénase de *Sphingomonas* CHY-1 (PhnI) dans *E. coli*.

A : Gel SDS PAGE de protéines montrant l'expression de PhnI dans *E. coli*. - Puits 1 : souche BL21(DE3)(pSD9)(pEB421), puits 2-4 : 3 clones DH5α(pFMA1)(pIBA34). L'expression a été induite par l'ajout de 1 mM d'IPTG

PhnA1 : Sous-unité α de PhnI - PhnA2 : Sous-unité β - PhnA4 : Réductase. La ferredoxine PhnA3 (12 kDa) n'est pas visible sur ce gel.

B : Dosages de l'activité des dioxygénases *in vivo* avec le phénanthrène ou l'anthracène comme substrats. L'histogramme compare l'activité de 3 clones DH5α(pFMA1)(pIBA34) avec celle mesurée dans la souche BL21(pSD9)(pEB431). Les cultures ont été incubées 24h en présence de chacun des HAP, puis les diols produits ont été extraits et dosés par GC/MS.

Afin d'obtenir des enzymes pleinement actives, le vecteur pFMA1 a ensuite été co-transformé avec pIBA34 portant *phnA3* et *phnA4* (les transporteurs d'électrons) dans la souche DH5 α . Les deux vecteurs possèdent des origines de répllication compatibles (ColE1/rep issu de pBRC1 respectivement) ainsi que des gènes de résistance aux antibiotiques différents (Ampicilline/Kanamicine) permettant leur maintien simultané dans les cellules hôtes. Des expériences comparant l'expression de PhnI dans DH5 α (pFMA1)(pIBA34) et dans le système BL21(DE3)(pSD9)(pEBA43) précédemment utilisé au laboratoire ont montré que la dioxygénase est presque aussi active dans le premier cas que dans le second (Figure IV.40).

3.4 Construction des dioxygénases hybrides dans *E. coli*

3.4.1 Clonage de domaines catalytiques de dcatdiox A dans pFMA1 par la méthode PIPE

Le principe du clonage PIPE est décrit dans matériels et méthodes. Pour amplifier un fragment correspondant au vecteur pFMA1, omettant le domaine catalytique de PhnI, les amorces PhnRHDr-V/PhnRHDf-V ont été utilisés (leur séquence et le programme d'amplification sont détaillés en annexes A). Une autre paire d'amorces, Phn-RHDf22 et Phn-RHDR32, a été spécifiquement conçue pour amplifier les séquences dcatdiox A en ajoutant aux extrémités, une quinzaine de paires de bases complémentaires des extrémités du vecteur (Figure IV.33).

Après l'amplification de pFMA1, le plasmide qui a servi de matrice à la PCR a été digéré par DpnI, une enzyme qui clive l'ADN méthylé (entre A et T). Cette étape permet de supprimer le vecteur matrice et donc de réduire le bruit de fond de clonage.

En ce qui concerne les séquences dcatdiox A, une première amplification (amorces Phn-RHDf2 et Phn-RHDR3) a été réalisée avec un temps d'élongation de 20 s. Elle a été suivie par seconde PCR (amorces Phn-RHDf22 et Phn-RHDR32) avec un temps d'élongation réduit à 14 s afin de favoriser la formation de simples brins aux extrémités des inserts (Figure III.3 de Matériel et Méthodes).

Les produits de PCR purifiés ont été mélangés dans un ratio 4 : 1 (insert : vecteur) avant transformation (cellules NEB 5 α High Efficiency).

Une dizaine de constructions de gènes hybrides a été tentée mais seulement deux hybrides ont été obtenus avec les séquences U3-111 et BB10 (Figure IV.36, encadrés violets). Les plasmides correspondant, pFMA8 et pFMA9, ont été transférés avec pIBA34 dans *E. coli* DH5 α . Le dosage de l'activité des dioxygénase *in vivo* a montré que ces hybrides n'étaient actifs ni sur le phénanthrène ni sur l'anthracène.

3.4.2 Clonage de séquences dcatdiox A plus courtes dans pFMA1

L'absence d'activité enzymatique des hybrides obtenus avec les séquences U3-111 et BB10 pourrait être la conséquence d'un mauvais repliement des protéines ou d'un défaut d'interaction entre le domaine Rieske et le domaine catalytique.

Nous avons donc décidé de réduire la taille de la séquence remplacée dans pFMA1, codant le domaine catalytique. Les séquences U3-111 et BB10 possèdent, à 323 pb de leur extrémité 5' un site unique EcoRV, présent aussi dans *phnA1a*. Ce site a servi à construire des gènes hybrides selon le schéma décrit à la Figure IV.41, dans les vecteurs pFMA10 (U3-111) et pFMA11 (BB10). Les constructions ont été vérifiées par séquençage.

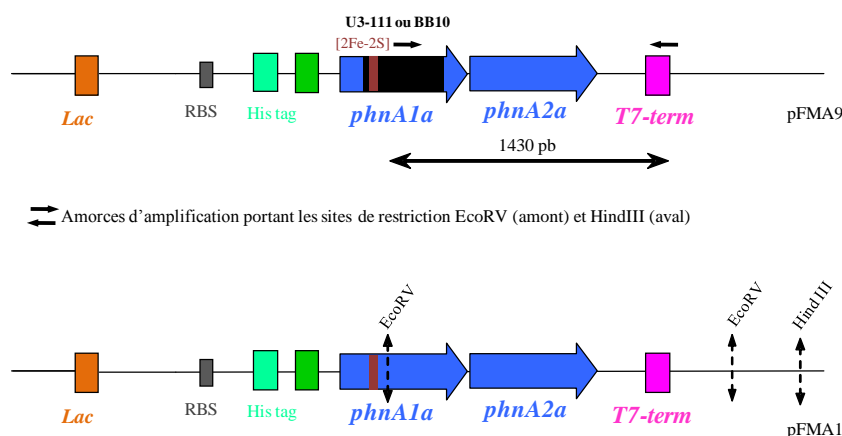


Figure IV.41 : Schéma de clonage des séquences U3-111 et de BB10 (624 pb) dans le vecteur pFMA1. Les amorces amont et aval amplifient un fragment de 1430 pb comprenant un site EcoRV en 5' et un site HindIII 5 en 3'. Le fragment a été cloné entre les sites EcoRV et HindIII de pFMA1.

L'expression de ces gènes hybrides dans *E.coli* DH5 α contenant pIBA34 a produit des dioxygénases inactives *in vivo* en présence de phénanthrène comme substrat.

L'absence d'activité des hybrides pourrait provenir d'une instabilité des protéines produites, du fait d'une trop grande différence de séquence entre U3-111, BB10 et PhnA1a. En effet, ces deux séquences ne possèdent que 60% et 58% d'identité avec la partie correspondante de PhnA1a.

Afin de tester cette hypothèse, six séquences représentatives du groupe dcatdiox A, ont été choisies pour construire de nouveaux hybrides : U3-16, U3-60, U3-72, U3-89, U3-116 et U3-124 (Figure IV.36 – encadrés verts). Leur pourcentage d'identité avec la séquence correspondante de PhnI varie de 82% (U3-116) à 36% (U3-16 et U3-124).

Les séquences U3-60, U3-72, U3-89 et U3-116 ont été amplifiées par PCR en utilisant des amorces introduisant des sites RsrII et AgeI aux extrémités, puis clonés entre les sites correspondant dans pFMA1 (entre les bases 545 et 1240 de *phnA1a*, soit entre Gly183 et Phe413 ; 0 et Figure IV.42). Les séquences U3-16 et U3-124 étant plus courtes que les autres (826 pb), nous n'avons pas pu utiliser le site AgeI. Nous avons donc cloné un fragment plus petit RsrII-MluI (entre les bases 545 et 1125, soit entre Gly183 et Ile375 ; 0 et Figure IV.42). Bien que plus court, le fragment substitué code pour une partie plus restreinte mais complète du domaine catalytique, portant notamment les trois ligands du site actif (0).

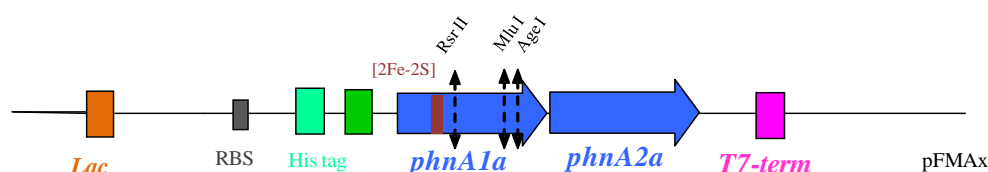


Figure IV.42 : Schéma de principe de la construction des hybrides contenant les domaines catalytiques U3-16, U3-60, U3-72, U3-89, U3-116 et U3-124.

Toutes les constructions ont été vérifiées par séquençage, pour s'assurer notamment que les jonctions de part et d'autre de l'insert se sont bien faites dans la phase de lecture du gène *phnA1a*.

3.5 Surexpression des dioxygénases hybrides dans *E. coli*

Les vecteurs obtenus, pFMA2 (U3-116), pFMA3 (U3-60), pFMA4 (U3-72), pFMA5 (U3-89), pFMA6 (U3-124) et pFMA7 (U3-16), ont été co-transformés avec pIBA34 dans *E. coli* DH5 α . Cependant, nous avons constaté que l'expression des protéines recombinantes, y compris PhnI, était instable dans cette souche.

Les plasmides pFMA2 à pFMA7 et pIBA34 ont donc été transférés dans *E. coli* JM109, une souche possédant notamment le répresseur de l'opéron lactose *lacI^q* qui permet de mieux contrôler l'expression (Yanisch-Perron *et al.*, 1985).

Après induction pendant 16h à 25°C, les extraits totaux des souches recombinantes ont été analysés par SDS-PAGE (Figure IV.43).

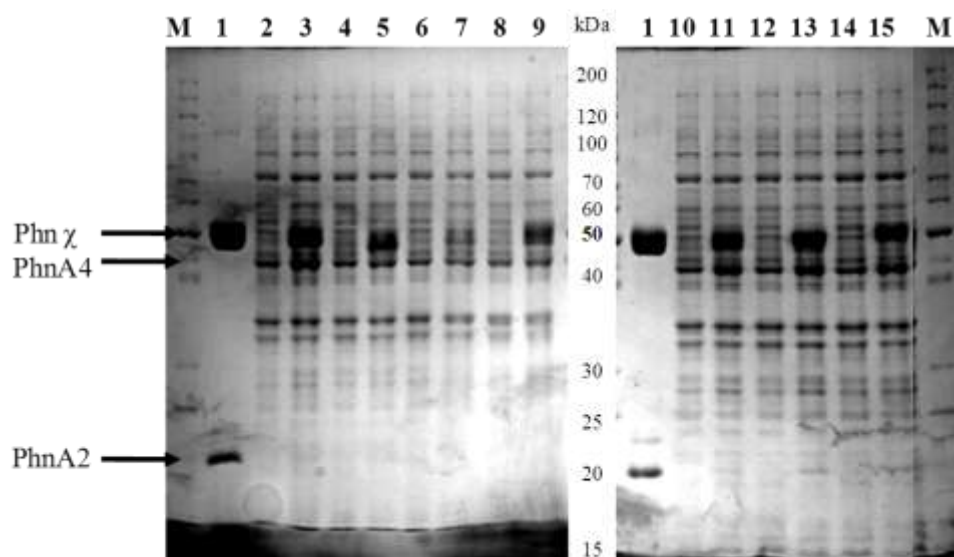


Figure IV.43 : Expression des protéines hybrides (Phn χ) dans *E.coli* JM109 (pIBA34). Pour chaque hybride, deux échantillons ont été prélevés avant induction (ti) et en fin d'incubation par 1 mM IPTG (tf). Les extraits protéiques totaux ont été déposés sur le gel par ordre décroissant de similarité des hybrides avec PhnI.

Puits 1 : PhnI purifiée ; Puits 2 : (pFMA1) ti ; Puits 3 : (pFMA1) tf ; Puits 4 : (pFMA2) ti ; Puits 5 : (pFMA2) tf ; Puits 6 : (pFMA3) ti ; Puits 7 : (pFMA3) tf ; Puits 8 : (pFMA4) ti ; Puits 9 : (pFMA4) tf ; Puits 10 : (pFMA5) ti ; Puits 11 : (pFMA5) tf ; Puits 12 : (pFMA6) ti ; Puits 13 : (pFMA6) tf ; Puits 14 : (pFMA7) ti ; Puits 15 : (pFMA7) tf

Le résultat montre que les protéines hybrides U3-116, U3-89, U3-124 et U3-16 sont exprimées à un niveau équivalent à PhnI, à en juger par l'intensité de la bande de 50 kDa (sous-unité alpha). En se basant sur ce critère, les hybrides U3-60 et U3-72 sont produits à un niveau sensiblement plus faible (Figure IV.43). La sous-unité beta est peu visible sur ce gel,

ce qui suggère qu'une partie au moins des sous-unités α n'est pas associée à une proportion équimolaire de sous-unités β .

3.6 Dosage de l'activité des dioxygénases hybrides et caractérisations préliminaires

L'activité des hybrides a ensuite été mesurée en présence de 4 HAP substrats : le naphthalène et le biphenyle (2 cycles), le phénanthrène et l'anthracène (3 cycles) (Figure IV.44)

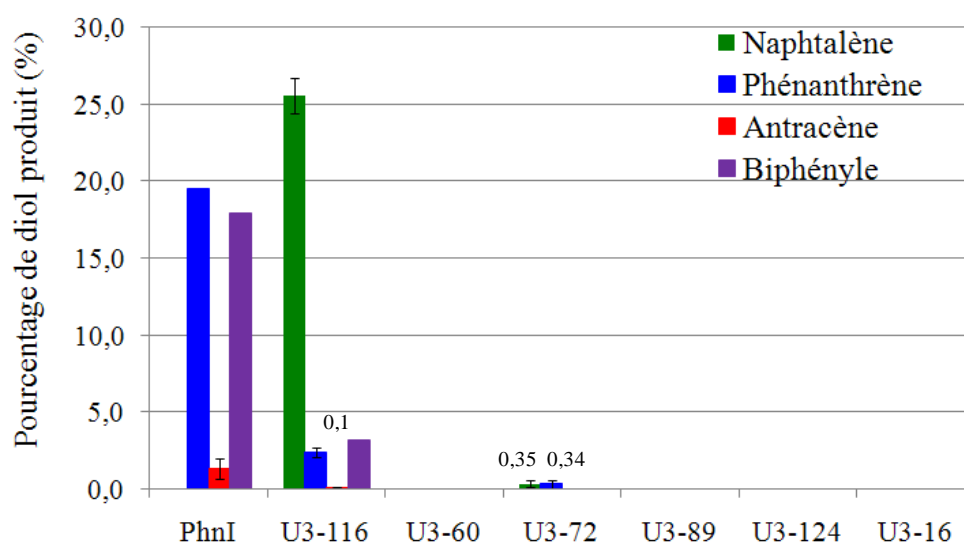


Figure IV.44 : Dosage de l'activité des dioxygénases hybrides sur quatre HAP substrats. Les résultats sont exprimés en pourcentage de l'activité maximale mesurée pour PhnI en présence de Naphtalène. Les hybrides sont classés dans le sens de la lecture par ordre décroissant de leur pourcentage de similarité avec la séquence de PhnI.

L'hybride U3-116, qui présente le pourcentage de similarité le plus élevé avec PhnI (82% sur le segment 183-413) est capable d'oxyder les 4 HAP. Comparée à l'activité de PhnI, l'activité de cet hybride est inférieure d'un facteur 4 à 5 sur les HAP de deux cycles et d'un facteur 8 à 13 sur les HAP de trois cycles. Cette observation suggère que l'hybride U3-116 est moins compétent que PhnI dans l'oxydation des HAP de deux et trois cycles.

Une activité faible mais significative a été mesurée pour l'hybride U3-72 en présence des substrats naphthalène et phénanthrène, alors que l'hybride U3-60 n'a pas montré d'activité détectable. Ces deux hybrides présentent 80% de similarité avec PhnI et 78% avec U3-116 (sur le segment 183-413). Cette faible activité peut en partie s'expliquer par un niveau d'expression relativement bas par rapport aux autres hybrides, comme l'a montré l'analyse

SDS-PAGE (Figure IV.43). D'ailleurs, l'hybride U3-72 dont le niveau d'expression est supérieur à celui de l'hybride U3-60 est aussi celui qui présente une activité détectable.

Ce résultat suggère que l'hybride U3-60 et dans une moindre mesure U3-72, sont instables, et donc produits en quantité limitée dans *E. coli*.

Enfin, aucune activité n'a été détectée pour les hybrides U3-89, U3-124 et U3-16 qui partagent 59%, 36% et 36% de similarité de séquence avec PhnI. L'expression des protéines hybrides est pourtant équivalente à celle de PhnI, à en juger par la bande à 50 kDa détectée dans l'extrait de protéines total. Cela laisse supposer que ces protéines hybrides sont soit inactives, soit synthétisées sous forme de corps d'inclusion insolubles.

3.7 Purification des dioxygénases hybrides

Les dioxygénases hybrides ont été exprimées avec, en N-terminal, une étiquette poly-histidine (Figure IV.39). Ce type d'étiquette est généralement utilisé pour la purification des protéines par chromatographie d'affinité (IMAC). Les protéines hybrides ont été exprimées dans la souche JM109 dans 800 ml de culture. Après lyse des bactéries et centrifugation haute vitesse, l'extrait protéique a été déposé sur une colonne d'affinité sur laquelle est fixé du cobalt (Talon). Chaque protéine hybride a été élué de la colonne en fractions de 1 ml par un tampon contenant 150 mM d'imidazole.

Les fractions purifiées ont été analysées par spectrophotométrie UV-visible entre 250 et 700 nm (Figure IV.45-A). La dioxygénase PhnI possède un spectre caractéristique avec trois bandes d'absorption à 320, 420 et 460 nm dues au centre [2Fe-2S] de type Rieske (tracé bleu). Un spectre comparable a été obtenu avec l'hybride U3-116 (tracé vert et insert), indiquant que cette protéine a bien intégré le centre [2Fe-2S]. Par contre, les autres préparations ne montrent pas ce spectre caractéristique.

L'analyse SDS PAGE des fractions purifiées montre que l'hybride U3-116 est bien composé de deux types de sous-unités de 50 et 20 kDa dans des proportions voisines de celles de la protéine témoin PhnI (Figure IV.45-B). Une bande de faible intensité à 50 kDa est aussi détectée dans le cas de l'hybride U3-72, confirmant qu'une petite partie seulement de cet hybride est produite sous forme soluble et active.

Dans les autres cas, l'analyse ne montre que des bandes de protéines contaminantes, suggérant qu'aucun des autres hybrides n'a été produit sous forme soluble. Cette hypothèse a été validée par une analyse Western Blot avec des anticorps dirigés contre l'étiquette polyhistidine des fractions solubles du lysat cellulaire (après centrifugation haute vitesse). Le résultat a montré qu'une bande à 50 kDa était détectable pour les lysats contenant PhnI et U3-116, mais aucune bande spécifique n'a été observée pour les lysats censés contenir les hybrides U3-60, U3-72, U3-89, U3-124 et U3-16.

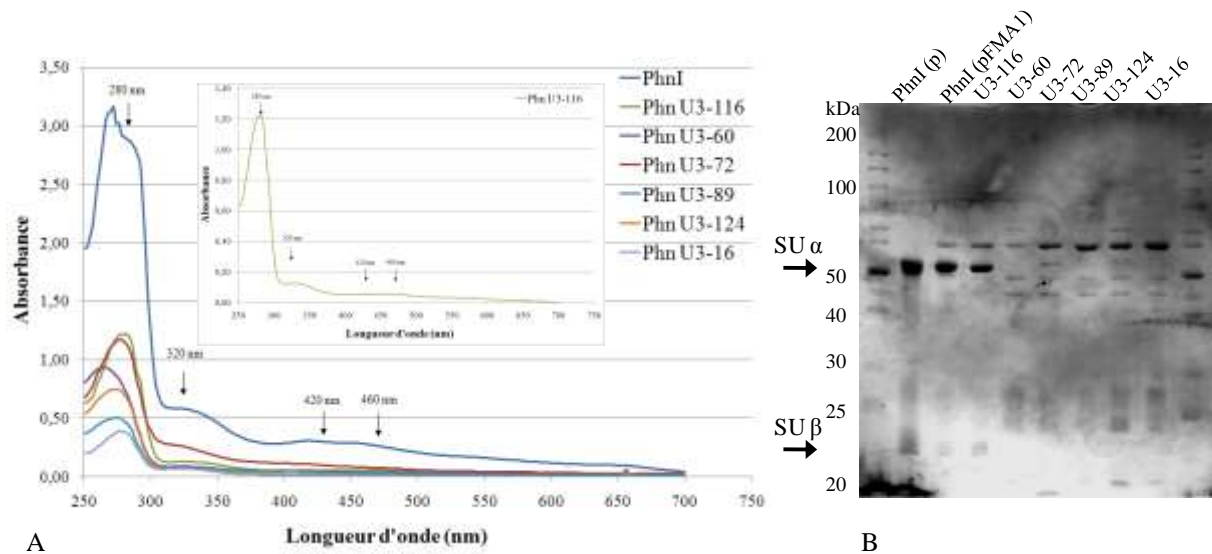


Figure IV.45 : Purification des dioxygénases hybrides et de PhnI par chromatographie d'affinité IMAC.
A : spectres d'absorbance des fractions purifiées. L'insert présente un zoom du spectre de U3-116.
B : Analyse SDS PAGE des fractions purifiées. Un échantillon de PhnI purifié au laboratoire (p) a servi d'étalon.

En conclusion, ces résultats indiquent que tous les hybrides, à l'exception d'U3-116, sont instables et ne sont pas produits en quantité appréciable dans *E. coli*. Cette instabilité se traduit soit par la formation de corps d'inclusion insolubles (U3-16, U3-89 et U3-124), soit par un défaut de production de la protéine recombinante (U3-60 et U3-72). Ces phénomènes pourraient s'expliquer par une trop grande dissemblance entre les domaines catalytiques de PhnI et celui provenant des enzymes issues du sol. Ceci pourrait entraîner un mauvais repliement de la sous-unité α des hybrides et/ou un mauvais assemblage de l'hexamère $\alpha_3\beta_3$.

3.8 Caractérisation de l'activité des domaines catalytiques de Burkholderiales issus des microcosmes

Dans la même optique que précédemment, nous avons voulu construire un système d'expression d'enzymes hybrides à partir des sous-unités α (PhnAc) et β (PhnAd) de la dioxygénase de *Acidovorax* sp. NA3 dans laquelle le domaine catalytique serait remplacé par des séquences du groupe dcattediox B.

Le gène *phnAcAd* a été amplifié à partir de l'ADN génomique de la souche NA3 à l'aide des amorces *phnAcNdeI/phnAdXhoI* qui ajoutent des sites *NdeI* et *XhoI* aux extrémités 5' et 3' de la séquence. L'amplicon (1916 pb) a été inséré dans le vecteur pJET1.2 par clonage bouts francs pour donner le vecteur pFMA12. Les gènes *phnAcAd* ont été sous-clonés par digestion *NdeI/XhoI* dans le vecteur pET15 pour donner le plasmide pFMA13 et la construction a été vérifiée par séquençage (Figure IV.46).

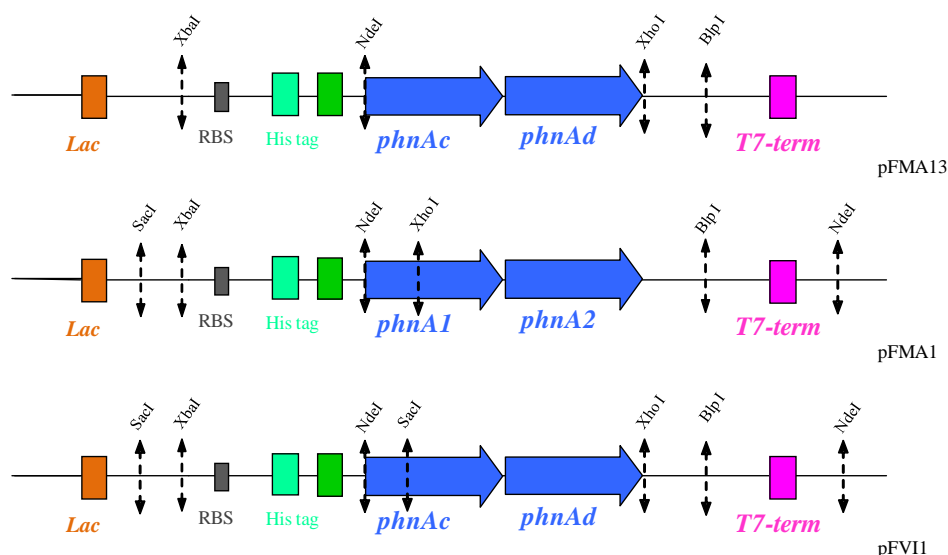


Figure IV.46 : Schéma du vecteur pFVII1, un dérivé de pUC18 dans lequel la séquence de *phnAcd*, codant la dioxygénase de *Acidovorax* sp. NA3, a été introduite.

Un fragment de 2072 pb contenant *phnAcd* ainsi que le site de fixation du ribosome (RBS) et l'étiquette poly-histidine a enfin été sous-cloné de pFMA13 par digestion *XbaI* et *BpII* dans pFMA1 pour donner le vecteur pFVII1.

Les vecteurs pFMA13 et pFVII1 ont ensuite été transférés respectivement dans BL21(DE3) et dans JM109, avec ou sans pIBA34 qui porte les gènes codant les transporteurs d'électrons de *Sphingomonas* CHY1 (*phnA3* et *phnA4*). Suite à 16 heures d'induction par 1 mM d'IPTG à

25°C, deux protéines de 50 kDa et 20,8 kDa ont été détectées dans l'extrait protéique de *E. coli*. Ces protéines ont bien la taille prédite des sous unités PhnAc et PhnAd de la dioxygénase de la souche NA3. (Figure IV.47-A). Sur la base de l'analyse SDS PAGE, l'expression est légèrement meilleure dans le système BL21 (DE3)(pFMA13) que dans le système JM109 (pVFI1).

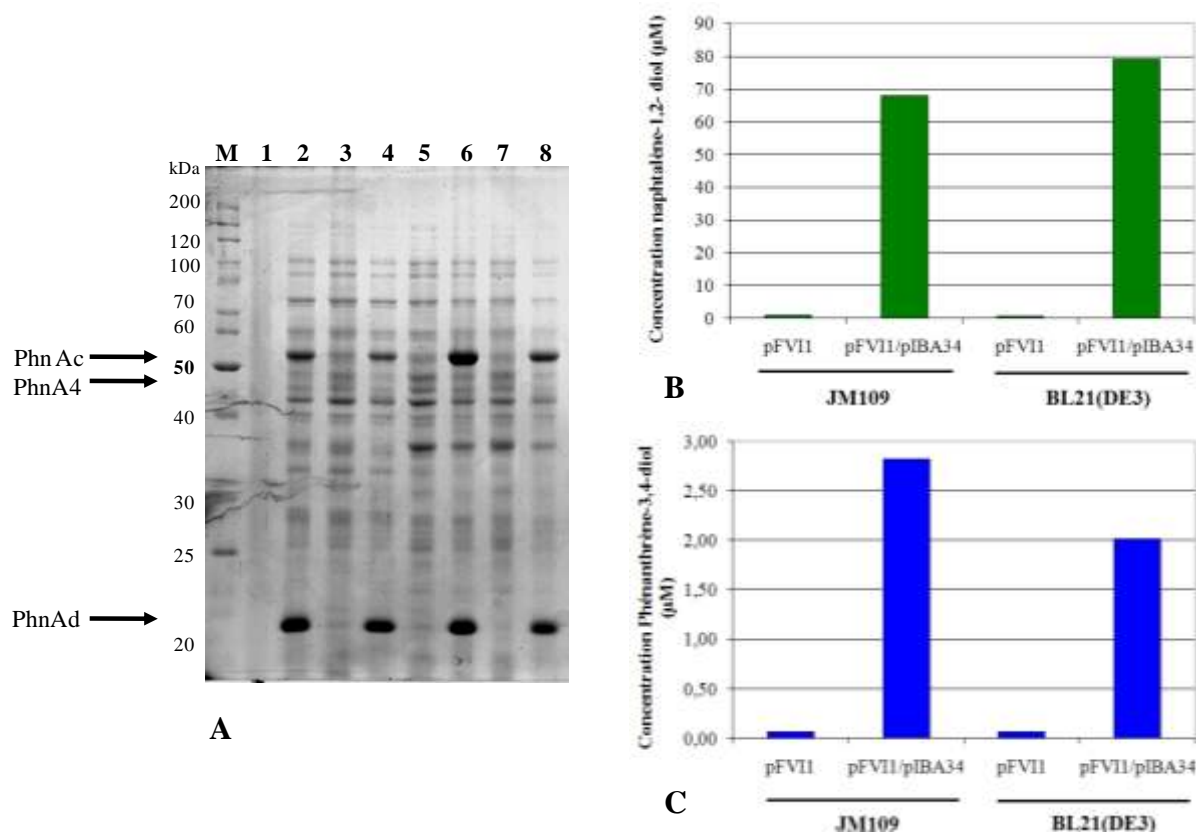


Figure IV.47 : Expression dans *E. coli* et dosage de la dioxygénase de *Acidovorax* sp. NA3.

A : expression de PhnAc et PhnAd dans JM109 en absence (puits 1 et 2) ou en présence (puits 3 et 4) des transporteurs d'électrons PhnA3 PhnA4 de *S. CHY-1* portés par pIBA34, et dans BL21(DE3) en absence (puits 5 et 6) ou en présence (puits 7 et 8) des transporteurs d'électrons portés par pIBA34. Pour chaque essai, un échantillon a été prélevé avant induction par 1 mM d'IPTG (puits 1, 3, 5, 7) et à la fin de l'expression (puits 2, 4, 6, 8).

B : dosage de l'activité de la dioxygénase de NA3 sur le naphthalène, en absence ou en présence des transporteurs d'électrons, dans JM109 ou BL21(DE3)

C : dosage de l'activité de la dioxygénase de NA3 sur le phénanthrène, en absence ou en présence des transporteurs d'électrons, dans JM109 ou BL21(DE3)

Le dosage de l'activité montre que la dioxygénase de la souche NA3 sous forme recombinante dégrade le naphthalène et le phénanthrène *in vivo* avec une préférence pour le naphthalène (Figure IV.47-B et C). De plus, la co-expression avec les transporteurs d'électrons PhnA4 et PhnA3 a eu pour effet d'augmenter l'activité de PhnAcAd d'un facteur 100 à 160 sur naphthalène et d'un facteur 30 à 50 sur phénanthrène (suivant le système d'expression). Ce

résultat montre que d'une part l'oxygénase a besoin de donneurs d'électrons spécialisés pour fonctionner de façon satisfaisante et que d'autre part, les transporteurs d'électrons de la souche CHY-1 sont compatibles avec l'oxygénase de la souche NA3, produisant une enzyme fonctionnelle.

Les deux systèmes d'expression ainsi créés peuvent être utilisés pour d'une part caractériser la dioxygénase de *Acidovorax* sp. NA3 et d'autre part étudier l'activité associée aux séquences du groupe dcatdiox B par la construction d'enzymes hybrides.

4. Comparaison des séquences peptidiques des dioxygénases hybrides, et modélisation de leur poche catalytique

Précédemment, la dioxygénase de *Sphingomonas* sp. CHY-1 a été caractérisée du point de vue structural (Jakoncic *et al.*, 2007b). Une vingtaine d'acides aminés bordant le site actif, ainsi que des motifs particuliers (boucles LI et LII) ont été mis en évidence (les détails de structure sont présentés au paragraphe I.3.2.2 de l'Etude bibliographique)

En s'appuyant sur ces éléments de structure, les séquences des hybrides ont été alignées avec celle de PhnI (Figure IV.48), et les structures des poches catalytiques ont été modélisées (modélisations réalisés par Jean Jakoncic au Brookhaven National Laboratory, National Synchrotron Light Source, Upton, NY, USA).

L'analyse montre qu'une grande partie des acides aminés impliqués dans les interfaces entre les sous-unités α , et entre les sous-unités α et β , est conservée.

Cependant, la méthionine (Met364) présente dans le domaine catalytique entre deux hélices est remplacée, dans les hybrides U3-124 et U3-16, par un tryptophane (Trp364). Selon la modélisation, ce remplacement séparerait les deux hélices, et pourrait modifier la configuration voisine du site catalytique. De plus, l'arginine (Arg348), impliquée dans l'interaction entre les sous-unités α et β , est remplacée, dans les hybrides U3-124 et U3-16, par une glutamine (Gln348) et par une histidine (His348) dans U3-89. Cela pourrait influencer la formation du complexe $\alpha_3\beta_3$.

La modélisation de la poche catalytique prédit que celle des hybrides U3-16 et U3-124 est plus volumineuse que celle des autres hybrides, dont la taille serait proche de celle de PhnI.

L'alignement des séquences d'acides aminés des hybrides avec celle de PhnI montre que la majorité des résidus impliqués dans la conformation et la spécificité de la poche catalytique sont conservés sauf pour les hybrides U3-16 et U3-124 (Figure IV.48).

Les résidus impliqués dans la sphère de coordination du Fe(II) catalytique (Rouge ; Asn200, His207, His212, Asp360) et le résidu probablement impliqué dans le transfert des électrons entre le centre Rieske d'une sous-unité α et le fer catalytique de la sous-unité adjacente (Asp204) sont conservés dans tous les hybrides.

Dans U3-16 et U3-124, trois résidus contribuant à la forme de la poche catalytique sont remplacés par des acides aminés portant le même caractère hydrophobe (Leu356/Phe356, Gly205/Ala205, Val208/Ala208) mais avec des encombrements stériques plus importants. Enfin, Thr308 est remplacé par un acide aminé hydrophobe, Ile308.

La spécificité de PhnI pour une large gamme d'HAP (Jouanneau *et al.*, 2006) pourrait dépendre de deux boucles (LI et LII), situées à l'entrée de la poche catalytique (Jakoncic *et al.*, 2007b). La boucle LI est composée des résidus 221 à 238 et la boucle LII, plus petite s'étend des résidus 253 à 265.

Alors que les résidus Leu223 et Leu226 de LI et Ile253 et Ile260 de LII (vert clair) sont conservés dans U3-116, U3-60, U3-72 et U3-89, ils sont remplacés par des résidus qui changent leur propriétés d'hydrophobicité ou de charge dans les hybrides U3-16 et U3-124.

Dans l'ensemble, les hybrides U3-116, U3-60, U3-72 et U3-89 ne présentent que très peu de variations au niveau des résidus bordant la poche catalytique. En revanche, les hybrides U3-16 et U3-124 présentent très peu de résidus conservés par rapport à PhnI au niveau du domaine catalytique.

	1	10	20	30	40	50	60
PhnA1-CHY1	MSGDTTLVDTVNASQSRQVFWDNRDVEDLEIERIFSRRAWMLGHKSLLPKPGDFITTYMAE						
Phn-U3-116	MSGDTTLVDTVNASQSRQVFWDNRDVEDLEIERIFSRRAWMLGHKSLLPKPGDFITTYMAE						
Phn-U3-60	MSGDTTLVDTVNASQSRQVFWDNRDVEDLEIERIFSRRAWMLGHKSLLPKPGDFITTYMAE						
Phn-U3-72	MSGDTTLVDTVNASQSRQVFWDNRDVEDLEIERIFSRRAWMLGHKSLLPKPGDFITTYMAE						
Phn-U3-89	MSGDTTLVDTVNASQSRQVFWDNRDVEDLEIERIFSRRAWMLGHKSLLPKPGDFITTYMAE						
Phn-U3-16	MSGDTTLVDTVNASQSRQVFWDNRDVEDLEIERIFSRRAWMLGHKSLLPKPGDFITTYMAE						
Phn-U3-124	MSGDTTLVDTVNASQSRQVFWDNRDVEDLEIERIFSRRAWMLGHKSLLPKPGDFITTYMAE						
	70	80	90	100	110	120	
PhnA1-CHY1	DKIILSHQSDGTTFRAFINSCTHRGNQICHADSGNAKAFVCNYHGWVYQDGS�VDVPLES						
Phn-U3-116	DKIILSHQSDGTTFRAFINSCTHRGNQICHADSGNAKAFVCNYHGWVYQDGS�VDVPLES						
Phn-U3-60	DKIILSHQSDGTTFRAFINSCTHRGNQICHADSGNAKAFVCNYHGWVYQDGS�VDVPLES						
Phn-U3-72	DKIILSHQSDGTTFRAFINSCTHRGNQICHADSGNAKAFVCNYHGWVYQDGS�VDVPLES						
Phn-U3-89	DKIILSHQSDGTTFRAFINSCTHRGNQICHADSGNAKAFVCNYHGWVYQDGS�VDVPLES						
Phn-U3-16	DKIILSHQSDGTTFRAFINSCTHRGNQICHADSGNAKAFVCNYHGWVYQDGS�VDVPLES						
Phn-U3-124	DKIILSHQSDGTTFRAFINSCTHRGNQICHADSGNAKAFVCNYHGWVYQDGS�VDVPLES						
	130	140	150	160	170	180	
PhnA1-CHY1	RCYHNKLDKQELAAKSVRVETKGFIFGCHDPEAPSLDYLGEFRFYLDTIWEGGGAGLE						
Phn-U3-116	RCYHNKLDKQELAAKSVRVETKGFIFGCHDPEAPSLDYLGEFRFYLDTIWEGGGAGLE						
Phn-U3-60	RCYHNKLDKQELAAKSVRVETKGFIFGCHDPEAPSLDYLGEFRFYLDTIWEGGGAGLE						
Phn-U3-72	RCYHNKLDKQELAAKSVRVETKGFIFGCHDPEAPSLDYLGEFRFYLDTIWEGGGAGLE						
Phn-U3-89	RCYHNKLDKQELAAKSVRVETKGFIFGCHDPEAPSLDYLGEFRFYLDTIWEGGGAGLE						
Phn-U3-16	RCYHNKLDKQELAAKSVRVETKGFIFGCHDPEAPSLDYLGEFRFYLDTIWEGGGAGLE						
Phn-U3-124	RCYHNKLDKQELAAKSVRVETKGFIFGCHDPEAPSLDYLGEFRFYLDTIWEGGGAGLE						
	190	200	210	220	230	240	
PhnA1-CHY1	LLGPPMKSLLHCNWKVPVENFVGDGYHVGWTHAAALQGIGGELAGLAGNRADIPFDDLGL						
Phn-U3-116	LLGPPMKSILACNWKVPVENFVGDGYHVGWTHAAALQGIGGELAGLAGNRADIPFDDLGL						
Phn-U3-60	LLGPPMKSILACNWKVPVENFVGDGYHVGWTHAAALQGIGGELAGLAGNRADIPFDDLGL						
Phn-U3-72	LLGPPMKSILACNWKVPVENFVGDGYHVGWTHAAALQGIGGELAGLAGNRADIPFDDLGL						
Phn-U3-89	LLGPPMKSILACNWKVPVENFVGDGYHVGWTHAAALQGIGGELAGLAGNRADIPFDDLGL						
Phn-U3-16	LLGPPMKSILACNWKVPVENFVGDGYHVGWTHAAALQGIGGELAGLAGNRADIPFDDLGL						
Phn-U3-124	LLGPPMKSILACNWKVPVENFVGDGYHVGWTHAAALQGIGGELAGLAGNRADIPFDDLGL						
	250	260	270	280	290	300	
PhnA1-CHY1	QFTTRHGHGFGVLDNAATAIHVKRDGYVKYLEETRGGIREKFGPERERLYVGHWNCSIFP						
Phn-U3-116	QFTTRHGHGFGVLDNAATAIHVKRDGYVKYLEETRGGIREKFGPERERLYVGHWNCSIFP						
Phn-U3-60	QFTTRHGHGFGVLDNAATAIHVKRDGYVKYLEETRGGIREKFGPERERLYVGHWNCSIFP						
Phn-U3-72	QFTTRHGHGFGVLDNAATAIHVKRDGYVKYLEETRGGIREKFGPERERLYVGHWNCSIFP						
Phn-U3-89	QFTTRHGHGFGVLDNAATAIHVKRDGYVKYLEETRGGIREKFGPERERLYVGHWNCSIFP						
Phn-U3-16	QFTTRHGHGFGVLDNAATAIHVKRDGYVKYLEETRGGIREKFGPERERLYVGHWNCSIFP						
Phn-U3-124	QFTTRHGHGFGVLDNAATAIHVKRDGYVKYLEETRGGIREKFGPERERLYVGHWNCSIFP						
	310	320	330	340	350	360	
PhnA1-CHY1	NCSFLYGTNTFKIWHPRGPHEIEVWYTMVPSDADPATKSAIQREATRSFGTAGTLESDD						
Phn-U3-116	NCSFLYGTNTFKIWHPRGPHEIEVWYTMVPSDADPATKSAIQREATRSFGTAGTLESDD						
Phn-U3-60	NCSFLYGTNTFKIWHPRGPHEIEVWYTMVPSDADPATKSAIQREATRSFGTAGTLESDD						
Phn-U3-72	NCSFLYGTNTFKIWHPRGPHEIEVWYTMVPSDADPATKSAIQREATRSFGTAGTLESDD						
Phn-U3-89	NCSFLYGTNTFKIWHPRGPHEIEVWYTMVPSDADPATKSAIQREATRSFGTAGTLESDD						
Phn-U3-16	NCSFLYGTNTFKIWHPRGPHEIEVWYTMVPSDADPATKSAIQREATRSFGTAGTLESDD						
Phn-U3-124	NCSFLYGTNTFKIWHPRGPHEIEVWYTMVPSDADPATKSAIQREATRSFGTAGTLESDD						
	370	380	390	400	410	420	
PhnA1-CHY1	GENMSSATYVNRGVITRDGMNSTMGVGYEGPHVPYPGIVGISFIFGETSYRGFYRFWKEM						
Phn-U3-116	GENMSSATYVNRGVITRDGMNSTMGVGYEGPHVPYPGIVGISFIFGETSYRGFYRFWKEM						
Phn-U3-60	GENMSSATYVNRGVITRDGMNSTMGVGYEGPHVPYPGIVGISFIFGETSYRGFYRFWKEM						
Phn-U3-72	GENMSSATYVNRGVITRDGMNSTMGVGYEGPHVPYPGIVGISFIFGETSYRGFYRFWKEM						
Phn-U3-89	GENMSSATYVNRGVITRDGMNSTMGVGYEGPHVPYPGIVGISFIFGETSYRGFYRFWKEM						
Phn-U3-16	GENMSSATYVNRGVITRDGMNSTMGVGYEGPHVPYPGIVGISFIFGETSYRGFYRFWKEM						
Phn-U3-124	GENMSSATYVNRGVITRDGMNSTMGVGYEGPHVPYPGIVGISFIFGETSYRGFYRFWKEM						
	430	440	450				
PhnA1-CHY1	IDAPDWASVKANDDNWDSVFTNRNFWNEKLNAAE						
Phn-U3-116	IDAPDWASVKANDDNWDSVFTNRNFWNEKLNAAE						
Phn-U3-60	IDAPDWASVKANDDNWDSVFTNRNFWNEKLNAAE						
Phn-U3-72	IDAPDWASVKANDDNWDSVFTNRNFWNEKLNAAE						
Phn-U3-89	IDAPDWASVKANDDNWDSVFTNRNFWNEKLNAAE						
Phn-U3-16	IDAPDWASVKANDDNWDSVFTNRNFWNEKLNAAE						
Phn-U3-124	IDAPDWASVKANDDNWDSVFTNRNFWNEKLNAAE						

Figure IV.48 : Alignement des sous-unités α des dioxygénases hybrides avec celle de PhnI. Les séquences ont été alignées avec ClustalX. La numérotation des résidus est celle de la séquence de PhnI. Les résidus inchangés par rapport à la séquence de PhnI sont surlignés en gris. Les acides aminés surlignés en rouge dans la partie N-terminale sont les ligands du centre [2Fe-2S].

Dans le domaine catalytique de PhnI, les codes couleur sont les suivants :

- rouge : ligands du centre Fe(II)
- bleu foncé : acides aminés de la région proximale de la poche catalytique
- bleu claire : acides aminés de la région centrale de la poche catalytique
- vert claire : acides aminés conservés des boucles LI et LII
- vert foncé : acides aminés impliqués dans l'interaction des sous-unités alpha-alpha et alpha-beta

Les résidus des séquences hybrides qui diffèrent par rapport à PhnI sont surlignées en jaune.

Les nombreuses variations constatées entre les séquences peptidiques de PhnI, U3-116, U3-60, U3-72 et U3-89 d'une part et celles de U3-16 et U3-124 d'autre part sont certainement dues au fait que ces deux derniers domaines catalytiques ont été apparentés à une salicylate hydroxylases et non à une HAP-dioxygénase.

5. Discussion et conclusions

Ce chapitre a eu pour objet de développer une nouvelle méthode d'analyse de la diversité phylogénétique et catalytique des HAP-dioxygénases présentes chez les bactéries du sol capables de dégrader le phénanthrène. L'approche a été divisée en deux volets : l'analyse de la diversité de dioxygénases de Protéobactéries basée sur l'emploi de paires d'oligonucléotides appropriés, et la construction de systèmes d'expression visant à doser l'activité enzymatique des dioxygénases du sol.

5.1 Analyse de la diversité phylogénétique des HAP-dioxygénases des Protéobactéries.

Le premier jeu d'amorces qui a été conçu s'est appuyé sur l'alignement de séquences de dioxygénases provenant d'Alpha-, Beta-, et Gamma- Protéobactéries. En majorité, les séquences détectées étaient apparentées aux dioxygénases de Sphingomonadaceae (Alphaprotéobactéries). Ce résultat montre que cette famille de bactéries, connues pour leur implication dans la dégradation des HAP, est présente dans les microcosmes SIP2000 ayant reçu une forte concentration de phénanthrène (337 ppm) alors qu'elle était plutôt minoritaire selon l'analyse des gènes d'ARNr16S. En outre, un tiers des séquences de dioxygénases détectées présente moins de 60% de similarité avec des séquences de dioxygénases connues. Cela montre que la population des Alphaprotéobactéries acteurs de la dégradation du

phénanthrène est diversifiée et que les séquences répertoriées dans les bases de données ne décrivent qu'une fraction incomplète de la diversité des dioxygénases.

Un second jeu d'amorces a permis de détecter des séquences plus spécifiques des dioxygénases de Burkholderiales. Pour la plupart, les domaines catalytiques amplifiés (57/65) à partir du sol présentent moins de 65% de similarité avec des dioxygénases connues. Ces séquences, qui présentent entre elles au moins 97% de similarité, appartiennent probablement à un groupe de Betaprotéobactéries abondant dans les microcosmes SIP2000 à 5 jours.

Ces résultats montrent que les approches phylogénétiques (gènes d'ARNr16S) et fonctionnelles (gènes de dioxygénases) sont complémentaires dans l'exploration de la diversité bactérienne associée à la biodégradation des HAP.

Cette étude avait l'ambition de cibler la diversité des HAP-dioxygénases la plus large possible. En revenant sur les conditions de PCR employées pour amplifier les séquences dcatdiox A, on se rend compte que la paire d'amorces utilisée se compose de 96 oligonucléotides pour l'hybridation en 5' (6 bases variables sur 24), et de 1152 oligonucléotides pour l'hybridation en 3' (9 bases variables sur 22). L'amplification à partir de l'ADN métagénomique de sol a conduit à la détection de 51 séquences de dioxygénases sur 96 clones (53%). Pour la paire d'amorces du groupe dcatdiox B, 8 oligonucléotides composait l'amorce 5' (3 bases variables sur 24), et 128 l'amorce 3' (7 bases variables sur 22). Cela conduit à l'identification de 65 séquences de dioxygénases sur 80 clones séquencés (81%).

Chaque paire d'amorces susceptible d'amplifier un domaine catalytique de dioxygénase est donc fortement diluée dans le mélange réactionnel de PCR. Il en résulte que la probabilité d'amplifier certaines dioxygénases plus rares diminue quand la complexité des amorces augmente. Ce phénomène de dilution, associé aux choix des bases variables (degré de dégénérescence) lors de la conception des amorces explique que certains groupes de dioxygénases, telles que celles issues des Gammaprotéobactéries notamment, n'ont pas été détectées.

Ainsi, pour sonder la diversité de HAP-dioxygénases, il faudrait utiliser plusieurs jeux d'amorces, avec des degrés de variation correspondant à des sous-familles d'enzymes. Sur

cette base, on pourrait non seulement cibler des séquences proches de gènes connus, mais aussi des groupes de gènes phylogénétiquement plus éloignés.

Les dioxygénases impliquées dans l'hydroxylation de composés aromatiques ont été particulièrement étudiées. Ainsi, plus de 200 gènes ont été identifiés, parmi lesquels, 45 sont des HAP-dioxygénases de bactéries Gram⁻. Un inventaire des amorces d'amplification conçues dans le but de détecter les dioxygénases par PCR a été réalisé récemment (Iwai *et al.*, 2011). Bon nombre de ces amorces ont une spécificité étroite car elles ont été conçues sur la base d'un alignement d'un nombre restreint de séquences. Ainsi, l'alignement de HAP-dioxygénases de souches de *Pseudomonas* a permis de cibler des gènes du type *nah* ou *ndo* (Wilson *et al.*, 1999). Avec la multiplication des gènes de HAP-dioxygénases identifiés, des amorces censées couvrir une diversité plus large ont été conçues.

En utilisant des amorces ciblant les HAP-dioxygénases de Beta- et Gamma Protéobactéries, des gènes de type *Pseudomonas* et des gènes non affiliés ont été trouvés dans deux sols pollués aux hydrocarbures (Bordenave *et al.*, 2008). Dans une autre étude, on a cherché à quantifier les gènes de HAP-dioxygénases de bactéries Gram⁻ dans des sols et sédiments pollués aux HAP. Des amorces ont été construites par alignement d'une partie du domaine catalytique (séquence comprise entre Asp204 et Ser305) de gènes de dioxygénases d'Alpha-, de Beta- et de Gamma- Protéobactéries. Les séquences obtenues se sont révélées très proches des dioxygénases ayant servi de modèles pour la conception des amorces (> 94% de similarité) (Cebron *et al.*, 2008). D'autre part, Ding et collaborateurs ont mis au point un jeu d'amorces ciblant les dioxygénases de bactéries Gram⁻ et Gram⁺, capable de détecter des gènes apparentés (> 97% similarité) à des HAP-dioxygénases de Mycobactéries, de *Burkholderia* et de *Novosphingobium*. Plusieurs gènes assimilés à des sous-unités α de dioxygénases ne présentaient que 72% à 90% de similarité avec les séquences contenues dans les bases de données (Ding *et al.*, 2010).

Pour ces quatre études, les amorces construites ciblaient des séquences de 300 à 450 pb, correspondant à des portions du domaine de fixation du centre Rieske et/ou des portions du domaine catalytique.

Enfin un jeu d'amorces ciblant des régions proches de celles visées dans notre étude a été employé, non pas pour étudier la diversité sur de l'ADN de sol mais pour détecter les gènes de dioxygénases dans des souches isolées de sédiments exposés aux HAP (Zhou *et al.*, 2006).

En conclusion, il n'existe pas d'amorces universelles capables d'amplifier toutes les familles de dioxygénases, ni même des HAP-dioxygénases, du fait de l'extrême diversité de séquences et de l'absence de régions suffisamment conservées. Chaque jeu d'amorce cible une partie de la diversité et apporte parfois sa contribution à la découverte de nouveaux gènes.

Les deux jeux d'amorces que nous avons proposés s'inscrivent dans cette logique, et ont effectivement permis de détecter des gènes codant de nouvelles dioxygénases, distantes de celles répertoriées dans les bases de données jusqu'à présent.

5.2 Caractérisation fonctionnelle des dioxygénases associées à la dégradation du phénanthrène

Dans ce chapitre, nous avons aussi mis au point un système d'expression pour doser l'activité des dioxygénases de *Sphingomonas* sp. CHY-1 et de *Acidovorax* sp. NA3 dans *E. coli*. Ce système a été utilisé pour produire des enzymes hybrides, où le domaine catalytique de S. CHY-1 a été remplacé par des séquences équivalentes issues du sol. Nous avons ensuite tenté de déterminer la spécificité de ces hybrides pour différents HAP.

Les résultats montrent que le principe fonctionne à condition que le domaine catalytique de remplacement partage au moins 80% de similarité (séquence en acides aminés) avec le domaine d'origine.

Concernant les constructions d'hybrides inactives, nous avons constaté que les protéines étaient produites sous forme insoluble. L'analyse des séquences protéiques des hybrides et la comparaison avec les données structurales de PhnI ont montré que, dans le cas des séquences les plus éloignées (< 60% de similarité), des variations de séquence pouvaient affecter la stabilité du complexe $\alpha_3\beta_3$.

Enfin, il est probable que la fonction de certaines enzymes, représentées par les séquences U3-16 et U3-124, soit différente de l'hydroxylation des HAP. En effet, la séquence la plus

proche s'apparente à des enzymes de type salicylate-hydroxylase, qui interviennent en aval dans la voie de dégradation des HAP (Jouanneau *et al.*, 2007; Pinyakong *et al.*, 2004)

Une approche basée sur la construction d'enzymes hybrides a été décrite en 2003, pour étudier des biphenyle-dioxygénases. Le domaine catalytique de la dioxygénase BphA1 de *Burkholderia* sp. LB400 (résidus 170 à 393 vs 183 à 413 dans notre étude) a été remplacé par les domaines catalytiques de biphenyle-dioxygénases de souches de *Pseudomonas*, *Ralstonia* ou *Rhodococcus*. Les domaines remplacés présentaient 71 à 95% de similarité en acides aminés, par rapport au domaine d'origine. Les constructions obtenues ont produit des enzymes fonctionnelles dont la spécificité pour les polychlorobiphényles était comparable à celle des enzymes modèles utilisées dans cette étude. Les auteurs ont conclu que la construction d'hybrides fonctionnels pouvait s'appliquer à l'analyse de la diversité, et dépendait de la compatibilité de structure entre le domaine inséré et le gène receveur (Kahl and Hofer, 2003).

En conclusion, nous pouvons dire que le concept d'enzymes hybrides fonctionne pour les HAP-dioxygénases présentant au moins 80% de similarité avec la protéine d'origine (PhnI vs U3-116). Pour surmonter les difficultés liées à l'instabilité des enzymes hybrides, on pourrait essayer d'autres systèmes d'expression comme *Pseudomonas putida* KT2442 par exemple, qui a fait ses preuves dans la surproduction de la dioxygénase de la souche CHY-1 sous une forme plus stable que celle obtenue dans *E. coli* (Jouanneau *et al.*, 2006). D'autre part, la surproduction sous forme recombinante active de la dioxygénase PhnAcAd de *Acidovorax* sp. NA3 ouvre la voie à l'analyse fonctionnelle de séquences apparentées issues du sol (dcardiox B ; 60% à 94% de similarité) par la construction d'hybrides.

Chapitre V. Approche protéomique ciblant les arène-dioxygénases des bactéries impliquées dans la dégradation des HAP

Nous avons vu dans les chapitres précédents que l'analyse détaillée de l'ADN métagénomique des bactéries du sol, couplée au marquage isotopique, est un bon moyen d'investigation non seulement des bactéries responsables de la dégradation des polluants mais aussi de certaines de leurs enzymes. Cependant, la détection de gènes biomarqueurs spécifiques de la dégradation ne renseigne pas sur l'activité métabolique car on ne sait pas si ces gènes sont exprimés *in situ*. De ce point de vue, la mise en évidence de protéines et notamment d'enzymes du métabolisme des HAP est un meilleur indicateur de l'activation du potentiel métabolique des communautés bactériennes impliquées. Nous avons donc essayé d'identifier ce type d'enzymes dans le sol par une approche protéomique en collaboration avec le laboratoire d'Etude de la Dynamique des Protéomes (CEA Grenoble).

Du fait du marquage isotopique employé dans notre étude, on pouvait s'attendre à ce que ces enzymes, elles aussi issues des voies de biosynthèse, incorporent le carbone marqué. Grâce à ce marquage, nous pensions repérer plus aisément les protéines d'intérêt, dans la mesure où leurs peptides marqués au ^{13}C devaient théoriquement se distinguer des peptides non marqués par l'analyse en spectrométrie de masse.

Nous avons choisi de nous intéresser aux arène-dioxygénases produites par les populations bactériennes qui dégradent spécifiquement le phénanthrène. Cette approche nécessite d'isoler les protéines du sol, une matrice complexe composée de multiples substances dont certaines peuvent interférer avec l'analyse des protéines : les particules fines et les matières humiques.

L'extraction des protéines à partir d'échantillons de sol peut être réalisée selon deux approches : une approche directe visant à lyser les bactéries *in situ* afin d'extraire les protéines du sol (Benndorf *et al.*, 2007), et une approche indirecte consistant à isoler les cellules bactériennes du sol avant de les lyser afin d'en extraire les protéines (Maron *et al.*, 2007).

L'approche directe concerne toutes les protéines présentes dans le sol sans distinction, y compris les protéines extracellulaires, et les protéines fongiques ou d'origine végétale. Notre objectif est de cibler essentiellement les enzymes bactériennes impliquées dans la dégradation

des HAP et plus spécifiquement les arène-dioxygénases représentant une infime fraction du pool de protéines.

Nous avons donc opté pour une approche indirecte en deux étapes décrite ci-dessous.

1. Mise au point d'un protocole d'extraction des cellules bactériennes du sol

1.1 Critères de choix du procédé d'extraction des cellules bactériennes

Les sols issus des bassins de rétention du site d'étude contiennent essentiellement des matières sédimentées et des matières organiques d'origine végétale. Les transformations physiques, chimiques ou microbiologiques de la matière organique produisent des substances telles que les acides humiques, les acides fulviques et l'humine (Peña-Méndez *et al.*, 2005), dont la structure et la composition sont très complexes et mal connues (Hosse and Wilkinson, 2001).

L'argile et les matières humiques hydrophobes génèrent une matrice complexe de particules et d'agrégats stables (Maron *et al.*, 2006; Peña-Méndez *et al.*, 2005) au sein desquels les bactéries viennent se loger ou s'adsorber par interaction de leur paroi avec les composants de cette matrice (Bakken, 1985; Lindahl, 1996; Mayr *et al.*, 1999).

L'étape critique de l'extraction des cellules bactériennes du sol consiste donc à casser ces agrégats et à décrocher les cellules des particules de sol (Robe *et al.*, 2003). Le procédé de dispersion doit si possible assurer une bonne efficacité d'extraction sans compromettre l'intégrité des cellules.

Différentes méthodes de traitement chimiques (Lindahl, 1996) ou mécanique (Bakken, 1985; Robe *et al.*, 2003) ont été proposées.

Après l'étape de dispersion, les cellules bactériennes peuvent être isolées par deux méthodes. Une alternance de centrifugations à basse et haut vitesse permet de sédimenter les particules, laissant les cellules dans le surnageant, puis de sédimenter les bactéries en éliminant les contaminants solubles (Maron *et al.*, 2006; Robe *et al.*, 2003). Une autre approche consiste à séparer les cellules bactériennes sur un gradient de densité de Nicodenz, de Percoll ou de

Metrizamide. Dans ce cas, les particules sédimentent dans le culot, et les bactéries flottent à la surface du gradient (Robe *et al.*, 2003).

1.2 Optimisation du procédé d'extraction

Afin de faciliter la mise au point et l'évaluation des procédés expérimentaux que nous avons testés, les échantillons de sol ont été mélangés, au début de chaque essai, avec 10 ml d'une suspension ($DO_{600} = 1$) de *Pseudomonas* sp KT 2442 (pSD9) portant les gènes de la dioxygénase PhnI de *Sphingomonas* sp. CHY-1 (Jouanneau *et al.*, 2006). La souche KT2442 a été choisie car c'est une bactérie du sol dont le génome est connu. La présence du plasmide pSD9 apporte un marqueur de résistance à la tétracycline utile pour déterminer le rendement d'extraction des bactéries. De plus, il porte les gènes d'une protéine connue de type dioxygénase, qui peut servir de marqueur pour l'analyse protéomique. .

Les échantillons de sol (10 g) ont été repris soit dans une solution saline ($H_2O.NaOH$ 0,8%) ou soit dans un tampon pyrophosphate de sodium (0,05 M pH 6,4). Le pyrophosphate de sodium est un composé qui agit sur la matière humique en altérant la structure des polymères permettant d'améliorer l'efficacité de l'extraction des cellules (Lindahl, 1996). Suivant le pH imposé au milieu il est possible de précipiter l'humine (milieu alcalin) ou les acides humiques (milieu acide). Les échantillons de sol ont donc aussi été soumis à des lavages successifs avec des tampons de pyrophosphate de sodium 0,1 M à pH neutre (6,7), acide(4) ou basique (9), permettant d'éliminer une partie de la matière organique brune.

La dispersion des agrégats et le décrochage des cellules des particules de sol ont été réalisés par traitement mécanique (ultrason, broyage) associés ou non à l'effet de surfactants (Tween 20, Triton x100, PolyVinyl Poly Pyrolidone ajoutés à 0,5% final). Les surfactants sont utilisés pour faciliter le décrochement des bactéries, car ils déstabilisent les interactions entre les parois bactériennes (lipo-polysaccharides) et les composants du sol.

Afin de préserver l'intégrité des cellules bactériennes, les surfactants n'ont été ajoutés que lors de certains essais de broyage. Les traitements aux ultrasons et le broyage ont été réalisés par cycles de 1 min de traitement suivie de 5 min de repos dans la glace. Cette alternance est préférée à un traitement long qui pouvait endommager les cellules (Bakken, 1985).

Les échantillons de sol ainsi traités ont été déposés sur un coussin de Nicodenz et centrifugés. Le halo de cellules formé à la surface du Nicodenz a été prélevé, lavé quatre fois soit dans une solution saline soit dans un des tampons pyrophosphate. Le culot de cellules a ensuite été conservé à -20°C avant l'étape de lyse.

1.3 Choix du protocole le plus adapté au sol étudié

Au cours de la mise au point de la méthode d'extraction des cellules bactériennes du sol, différentes associations des traitements présentés ci-dessus ont été testées. Les lavages en pyrophosphate de sodium ont permis d'éliminer une partie des matières humiques, mais la fraction restante a rendu le halo bactérien indiscernable des contaminants après la centrifugation. Les traitements mécaniques ont donc préférentiellement été réalisés dans une solution saline (NaCl à 0,8%) : 30 s d'exposition aux ultrasons, suivies de 3 cycles de 1 min de broyage.

L'association du traitement aux ultrasons puis du broyage en présence de Tween20 a donné un résultat plus satisfaisant (Figure IV.49-B) que le traitement aux ultrasons seul (Figure IV.49-A). En effet, le halo de cellules contient moins de particules dans le cas B que dans le cas A.

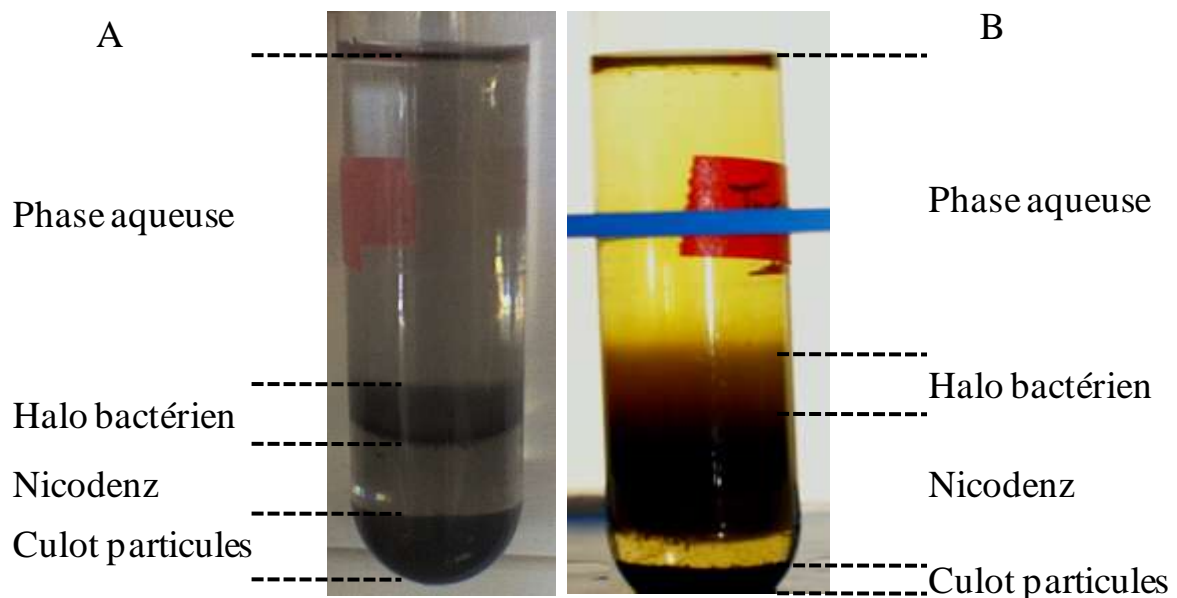


Figure IV.49 : Photos de deux tubes d'ultracentrifugation sur gradient de Nicodenz. A : échantillon de sol traité par ultrasons dans la solution saline (H₂O.NaCl 0,8%) puis déposé sur le coussin de Nicodenz. B : échantillon de sol traité par ultrasons dans la solution saline (H₂O.NaCl 0,8%), puis par broyage en présence de Tween20 avant ultracentrifugation sur gradient de Nicodenz.

Malgré les traitements relativement drastiques imposés aux échantillons de sol et le choix des conditions les plus favorables à la séparation des bactéries, les particules très fines n'ont pu être complètement éliminées des cellules bactériennes isolées sur gradient de Nicodenz. De plus, un suivi de la population de *Pseudomonas* sp. KT2442 par dénombrement sur boîtes de Pétri a montré qu'une proportion importante était perdue après centrifugation sur le gradient de Nicodenz : 1.10^6 ufc/ml pour $6,5.10^6$ ufc/ml ajouté au sol, soit 85% de perte.

Il est probable qu'une partie des cellules ait été entraînée dans le culot de particules, ou piégée dans les particules en suspension dans le Nicodenz.

2. Mise au point d'un protocole d'isolement des protéines des bactéries du sol

Dans la littérature, très peu d'études portent sur l'extraction de protéines à partir de sols. Les approches indirectes se sont souvent heurtées au problème de la contamination par des substances humiques ou phénoliques, qui gênent l'analyse protéique (Maron *et al.*, 2007).

Lors des essais d'isolement des protéines bactériennes à partir des bactéries du sol, un traitement mécanique par ultrasons a été employé. Quatre variantes ont été testées : (1) les cellules ont été reprises dans un tampon TE seul, (2) en présence de lysonase, un mélange de lysozyme et de nucléase, (3) de benzonase, (4) ou dans un tampon de lyse Urée-CHAPS-Triton-DTT. Quel que soit le traitement, les particules fines et des divers contaminants organiques ont empêché la détection des protéines par SDS PAGE.

Ce résultat pouvait s'expliquer soit par une mauvaise lyse cellulaire, soit par l'adsorption des protéines aux particules ou aux substances organiques, soit par une quantité insuffisante de bactéries.

Par la suite, une méthode inspirée de l'approche par extraction directe des protéines du sol a été envisagée (Benndorf *et al.*, 2007). Les bactéries isolées sur gradient de Nicodenz ont été reprises dans une solution de soude (0,2 M) et incubées 1 h à 90°C afin de lyser les cellules. Après neutralisation de la suspension en tampon MES, les protéines ont été extraites par du phénol dans un bain à ultrasons. Cette procédure permet de solubiliser les protéines préférentiellement dans la phase organique, tout en favorisant leur séparation des particules de

sol. Enfin, les protéines ont été purifiées soit par précipitation à froid en présence d'acétate d'ammonium dans le méthanol, soit par extraction du phénol par le diisopropyléther.

La précipitation à froid a donné les meilleurs résultats, car l'analyse par SDS PAGE de la préparation a révélé la présence de protéines, la plupart appartenant à la souche KT2442 utilisée comme traceur (Figure IV.50 puits 2).

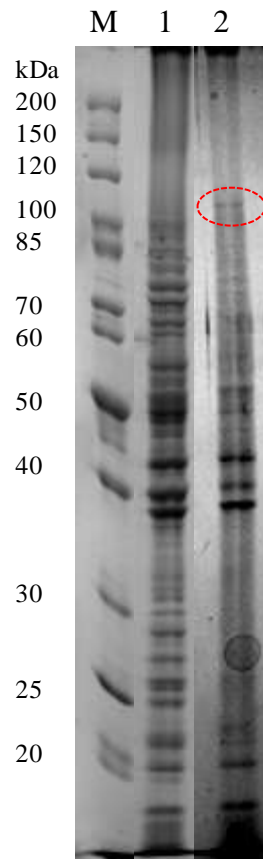


Figure IV.50 : Analyse SDS-PAGE des protéines extraites des bactéries du sol. Le halo bactérien du tube B de la Erreur ! Source du renvoi introuvable. **a été traité aux ultrasons dans un tampon TE. Les protéines ont été extraites par le phénol puis précipitées par l'acétate d'ammonium dans le méthanol.**

Puits 1 : extrait de *Pseudomonas* sp. KT2442

Puits 2 : échantillon bactérien isolé du sol (Figure IV.49, tube B).

Le cercle rouge localise une bande de protéines qui ne semble pas appartenir à *Pseudomonas* sp. KT2442.

Toutefois, une bande de taille proche de 100 kDa pourrait provenir d'une autre bactérie car on ne la trouve pas dans l'extrait de la souche KT2442 (entourée en rouge). Les protéines du sol sont minoritaires, indiquant que l'efficacité d'extraction est faible, sans doute à cause du faible rendement d'isolement des bactéries du sol.

3. Identification des protéines extraites du sol par analyse protéomique

La détection de protéines du sol par analyse protéomique n'a été que très peu documentée (von Bergen *et al.*, 2009), car il n'existe pas de méthode d'isolement garantissant un rendement d'extraction et une qualité de préparation suffisants (Jain and Chauhan, 2010). Cependant, l'étude des enzymes impliquées dans une fonction métabolique comme la dégradation de polluants est le moyen le plus direct de connaître la nature et le rôle de ces protéines au sein des communautés bactériennes (Jehmlich *et al.*, 2008).

Les protéines issues d'un échantillon environnemental (métaprotéome) constituent un ensemble extrêmement complexe vu la diversité d'espèces par gramme de sol (10^4 espèces estimées : voir le paragraphe I.1.1 de l'Etude bibliographique), le nombre moyen de protéines par bactérie (600 à 6000) (Jain and Chauhan, 2010) et le niveau variable de leur expression. Il est donc nécessaire d'effectuer un tri préalable afin de simplifier l'analyse protéomique.

Le marquage isotopique représente une approche séduisante qui devrait permettre de détecter préférentiellement les protéines des microorganismes ayant incorporé l'isotope ^{13}C du substrat (Jehmlich *et al.*, 2009). Par ce moyen, on pouvait espérer accéder plus facilement aux enzymes de la fonction métabolique ciblée afin de les identifier et de les quantifier (Snijders *et al.*, 2005).

Nous avons donc tenté de mettre en place une méthode d'analyse des protéines issues des microcosmes SIP, spécifiquement impliquées dans le métabolisme des HAP. Dans cette optique, et en préalable à l'analyse des protéines du sol, nous avons testé l'approche protéomique sur des protéines marquées issues de souches pures. Les analyses par spectrométrie de masse ont été réalisées au Laboratoire d'Etude de la Dynamique des Protéomes (CEA) en collaboration avec Yohann Couté.

3.1 Identification d'enzymes du métabolisme des HAP de souches pures cultivées sur phénanthrène marqué

Les souches *Sphingomonas* sp. CHY-1 et *Mycobacterium* sp. 6PY1 ont été cultivées au laboratoire en milieu liquide en présence de phénanthrène ^{13}C ou ^{12}C comme source unique de carbone et d'énergie afin d'induire l'expression des enzymes de dégradation y compris les

dioxygénases. Les protéines ont été extraites puis séparées par SDS PAGE. Dans le cas de *Mycobacterium* sp. 6PY1, les protéines de tailles voisines de 50 kDa ont été digérées par la trypsine puis soumises à l'analyse protéomique (exemple de *Mycobacterium* sp. 6PY1 Figure IV.51 ; puits 1 et 2). Les protéines de *Sphingomonas* sp. CHY1 ont été analysées dans leur ensemble, après migration dans le gel de concentration uniquement (non représenté sur la Figure IV.51).

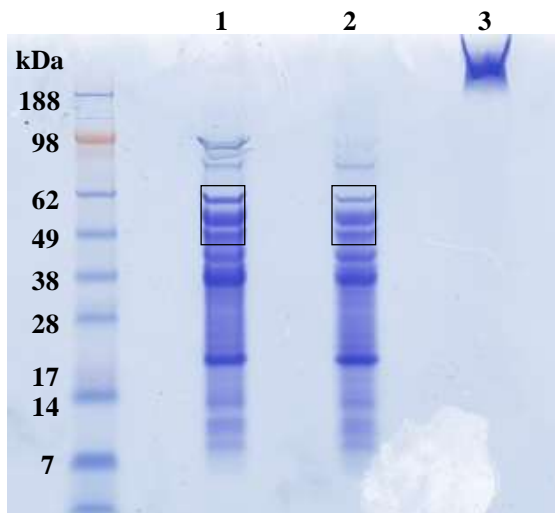


Figure IV.51 : Analyse SDS PAGE réalisée à la plateforme de protéomique, avant l'analyse protéomique

Puits 1 : protéines extraites de la souche pure *Mycobacterium* sp. 6PY1 cultivée sur phénanthrène non marqué

Puits 2 : protéines extraites de la souche pure *Mycobacterium* sp. 6PY1 cultivée sur phénanthrène marqué

Pour ces deux échantillons, les protéines situées dans le cadre ont été analysées par spectrométrie de masse

Puits 3 : protéines de l'échantillon de sol décrit Erreur ! Source du renvoi introuvable., puits 2.

Pour cet échantillon, l'électrophorèse a été stoppée une fois que l'échantillon est entré dans le gel de concentration pour une analyse globale de l'extrait de protéines.

L'analyse protéomique réalisée à la plateforme protéomique du laboratoire EDyP a permis de détecter les protéines des deux souches pures cultivées sur phénanthrène non marqué, et notamment les sous-unités alpha des dioxygénases. Ainsi, 132 protéines ont été détectées pour *Mycobacterium* sp. 6PY1 et 136 pour *Sphingomonas* sp. CHY1.

En revanche, les protéines de l'extrait de *Mycobacterium* marqué au ^{13}C n'ont pas été identifiées. Le spectromètre de masse n'a pas reconnu les peptides marqués et n'a donc pas converti les données en séquences peptidiques.

3.2 Détection des protéines issues d'échantillons de sol

L'échantillon de protéines du sol obtenu précédemment (paragraphe 2) a été déposé sur gel d'électrophorèse, mais sans le séparer Figure IV.51 puits 3, pour faire l'analyse protéomique de toutes les protéines de l'extrait.

Lors du traitement des données, les masses des peptides ont été comparées aux masses théoriques obtenues par digestion *in silico* des protéines des bases de données. La détection d'au moins deux peptides différents appartenant à la même protéine est jugée nécessaire pour s'assurer de la présence de cette protéine dans l'échantillon.

Les protéines (117 au total) ont été majoritairement reconnues comme appartenant au genre *Pseudomonas* (Figure IV.52), avec 30% d'entre elles appartenant à la souche *Pseudomonas* sp. KT2442 ajoutée dans le sol avant traitement. Parmi ces protéines, la sous-unité α de la dioxygénase de *Sphingomonas* sp. CHY-1 exprimée grâce au vecteur pSD9 ainsi que le répresseur LacI de l'opéron lactose de *E.coli* porté par ce vecteur, ont été détectés (8 peptides et 13 peptides). Les protéines identifiées et affiliées à *Pseudomonas* peuvent être classées en deux groupes :

- les protéines nécessaires au fonctionnement cellulaire (chaperonne, facteur d'élongation, ATP synthase, protéines de la membrane externe)
- les enzymes métaboliques avec une majorité de déshydrogénases (alcool déshydrogénase, aldéhyde déshydrogénase, acetyl-CoA déshydrogénase par exemple).

Ce sont des protéines abondantes dans les cellules bactériennes et non spécifiques d'une fonction particulière.

Certaines protéines ne sont pas directement associées au genre *Pseudomonas* (soit 17 protéines) :

12 protéines sont dites « litigieuses » (Figure IV.52) :

- la majorité des peptides qui ont servi à les identifier sont aussi présents dans des protéines de *Pseudomonas*. Les peptides uniques qui n'ont pas de correspondance avec les protéines de *Pseudomonas* (et qui font que ces protéines sont affiliées à d'autres genres bactériens) pourraient être des iso-formes des protéines de *Pseudomonas*, non encore connus ou caractérisés. Ce sont exclusivement des protéines indispensables au fonctionnement cellulaire (house keeping proteins).
- 1 protéine possède deux peptides sur sept non associés aux protéines de *Pseudomonas*. Théoriquement, les deux peptides nécessaires pour identifier cette protéine sont présents. Cette protéine chaperonne (60 kDa) pourrait donc être affiliée au genre *Actinobacillus* (Gammaprotéobactérie). Cependant, le fait que les cinq autres peptides soient aussi présents chez *Pseudomonas* laisse supposer que cette chaperonne est une

isoforme de protéines de ce dernier genre bactérien, et donc qu'elle ne provient pas forcément du sol.

- 2 protéines sont représentées par deux peptides sur trois non associés aux protéines de *Pseudomonas*. Il s'agit de chaperonnes affiliées aux genres *Thioalkalivibrio* (Gammaprotéobactérie) et *Acidovorax* (Betaprotéobactérie).
- 2 protéines sont représentées par des peptides que l'on ne retrouve pas dans *Pseudomonas*. Ces deux protéines chaperonnes sont affiliées à des Actinobactéries (*Mycobacterium* et *Corynebacterium*).

Les protéines non affiliées à *Pseudomonas* sont exclusivement des chaperonnes, des familles de protéines très abondantes dans les cellules bactériennes.

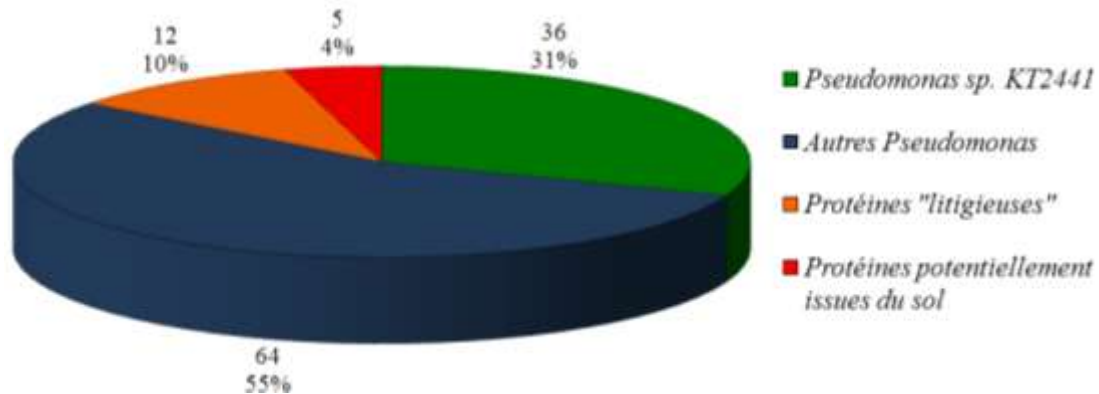


Figure IV.52 : Répartition des protéines identifiées par analyse protéomique. Les étiquettes pour chaque portion indiquent le nombre de protéines et la proportion de chaque groupe dans l'ensemble des protéines analysées. Sur les 36 protéines appartenant à *Pseudomonas sp. KT2442* une correspond à la sous-unité alpha de la dioxigénase de *Sphingomonas sp. CHY-1* (gène porté pas le plasmide pSD9 présent dans *P. KT2442*) et une autre au répresseur LacI de l'opéron lactose de *E.coli*, aussi présent sur pSD9. Le détail des affiliations des protéines est présenté en Annexe 6.

Au maximum 4% des protéines détectées sont potentiellement issues des microorganismes du sol. Ce faible rendement peut s'expliquer par les pertes lors des étapes d'extraction/purification, et notamment celles précédant la centrifugation sur gradient de Nicodenz. De plus, des matières organiques et des particules de sol pourraient être responsables de la séquestration des bactéries.

Il est aussi intéressant de noter que parmi les quelques protéines potentiellement issues du sol, on retrouve aussi bien des Protéobactéries (Gram⁻) que des Actinobactéries (Gram⁺) plus difficiles à lyser.

4. Discussion et conclusions

4.1 Extraction et purification des protéines à partir du sol

Les données présentées ci-dessus montrent que l'efficacité d'extraction du sol des cellules de la souche exogène (*Pseudomonas* sp KT2441) ne dépasse pas 15%. Cette souche, ajoutée au début du protocole d'extraction, n'est pas *a priori* affectée par les phénomènes d'absorption aux particules puisque son temps de séjour dans le sol est très court. Cependant la bactérie est plus exposée aux traitements chimiques et mécaniques, ce qui peut affecter sa viabilité. Le nombre de cellules isolées sur gradient Nycodenz est peut-être sous-estimé par le dénombrement de colonies sur boîtes de Pétri. Pour ces raisons, on ne peut pas conclure que l'efficacité de l'extraction des bactéries telluriques est comparable. Lors d'études comparables, l'efficacité d'extraction des bactéries du sol a été estimée par microscopie entre 10 et 25% seulement (Bakken, 1985; Courtois *et al.*, 2001; Lindahl, 1996; Maron *et al.*, 2006). Le rendement des cellules viables et cultivables ne dépasse pas 10% des cellules dénombrées avant extraction (Courtois *et al.*, 2001).

Les rendements sont d'autant plus faibles que les sols sont riches en argile et en matières organiques (Bakken, 1985; Mayr *et al.*, 1999). Ils pourraient être améliorés en appliquant plusieurs cycles de broyage/centrifugation sur les culots de sol, en utilisant des agents complexant tels que les cations divalents (Ca^{2+} apporté par une solution de CaCl_2) ou en augmentant la force ionique des solutions de lavage (Benndorf *et al.*, 2007; Criquet *et al.*, 2002; Lindahl, 1996).

Jusqu'à présent, l'extraction des cellules du sol était majoritairement destinée à l'extraction des acides nucléiques pour des analyses métagénomiques (Courtois *et al.*, 2001; Delmont *et al.*, 2011; Maron *et al.*, 2006). L'obtention de protéines du sol est nettement plus délicate notamment à cause de leur nature hydrophobe, surtout quand elles sont dénaturées, et du fait de leurs propriétés semblables à celles de la matière organique, lors des étapes d'extraction. Ainsi, les particules hydrophobes peuvent entraîner les protéines avec elles lors de précipitations ou de centrifugations. En outre, la présence d'acide humiques diminue la qualité des extraits protéiques (Singleton *et al.*, 2003) et peut fausser la quantification par les méthodes classiques (Benndorf *et al.*, 2007). L'association d'un traitement à la soude suivi

d'une extraction au phénol se traduit par un rendement peu satisfaisant (1%) lorsqu'il est appliqué à un sol brut (Benndorf *et al.*, 2007).

Cependant, la combinaison d'une séparation des bactéries sur gradient de Nicodenz avec une extraction préférentielle des protéines par la soude puis le phénol, nous a permis d'augmenter la qualité et la quantité de protéines extraites du sol et analysables par SDS PAGE.

Pour optimiser le protocole, il faudra encore améliorer le rendement d'extraction des bactéries en travaillant sur les étapes de dispersion et de centrifugation sur gradient.

4.2 Détection des protéines marquées par spectrométrie de masse

Au début de ce projet, nous souhaitions développer une méthode de détection des protéines spécifiquement induites par la présence d'un HAP, basée sur un marquage ^{13}C des protéines associée à la détection des protéines marquées, par protéomique. Dans un premier temps nous avons tenté de détecter des protéines de souches pures métaboliquement marquées avec le ^{13}C , mais le spectromètre de masse employé n'était pas paramétré pour cela.

Un peptide non marqué est composé à 99% de carbone ^{12}C et à 1% de ^{13}C (abondance naturelle de l'isotope dans l'environnement) alors que dans un peptide obtenu par marquage métabolique, le rapport $^{13}\text{C}/^{12}\text{C}$ est bien supérieur et dépend du taux d'incorporation de l'isotope. Au niveau du spectre de masse, l'isotope ^{13}C du carbone induit un changement du profil isotopique dépendant du pourcentage d'incorporation de l'isotope (Jehmlich *et al.*, 2008). Ainsi, le pic le plus abondant du profil est celui du peptide ayant incorporé la plus de ^{13}C . Le marquage n'étant pas homogène et total, des pics secondaires apparaissent. Non seulement le profil isotopique s'inverse mais en plus il s'élargit (Figure IV.53).

C'est pourquoi les profils des peptides marqués ont été jugés anormaux par le spectromètre de masse qui ne les a pas reconnus car il n'était pas paramétré pour cela.

Un paramétrage adéquat du spectromètre de masse rendrait possible la détection des peptides ayant un taux élevé d'incorporation du ^{13}C . Il subsisterait néanmoins le problème de l'identification des peptides, qui est réalisée par comparaison avec des profils créés *in silico* à partir des protéines présentes dans les bases de données.

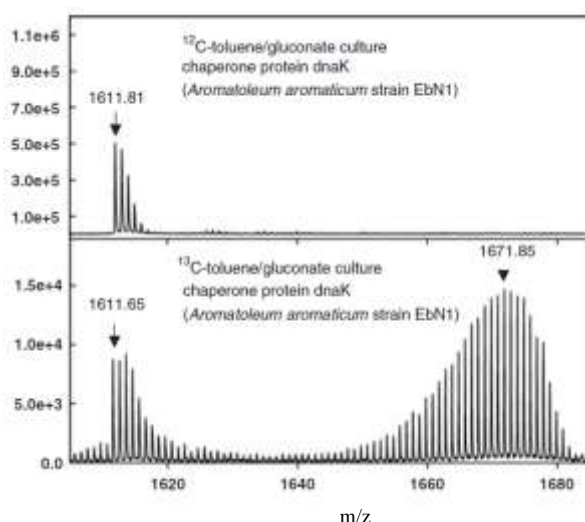


Figure IV.53 : Zoom sur le profil isotopique d'un peptide provenant d'une souche pure cultivée sur substrat non marqué ou marqué au ^{13}C (Jehmlich *et al.*, 2008).

Des études protéine-SIP publiées récemment apportent la preuve que ce type d'analyse est possible. Pour cela, des souches pures ont été cultivées en parallèle avec un substrat soit marqué soit non marqué. Des enzymes participant au métabolisme de ce substrat ont ensuite été détectées dans l'extrait protéique ^{12}C , puis le pourcentage d'incorporation de l'isotope lourd a été calculé dans le peptide marqué (Jehmlich *et al.*, 2008).

Une autre approche consiste à obtenir trois formes de protéines – non marquées ou marquées par les isotopes ^{13}C ou ^{15}N – dont les profils isotopiques sont comparés. L'identification se fait alors en calculant le nombre d'atomes de carbone et d'azote qui participent à la variation du profil isotopique. Cette technique permet d'appliquer des contraintes d'identification plus élevées (composition atomique), diminuant le nombre de correspondances possibles avec les protéines issues des bases de données (Snijders *et al.*, 2005).

Dans ces deux exemples, la preuve de concept de l'identification de protéines marquées a été apportée grâce à un traitement de données *a posteriori*. L'identification a été réalisée sur des peptides non marqués comparés aux bases de données. Puis sur ces résultats, les spectres de peptides marqués ont été détectés. En l'état actuel, la technique protéine-SIP ne permet donc pas encore d'identifier *a priori* des protéines marquées inconnues, directement à partir de leur spectre de masse.

De ce fait, la détection et l'identification de protéines marquées, dans le cas d'un échantillon très complexe (tel qu'un extrait protéique du sol) composé d'un petit nombre de peptides marqués, n'est pour le moment pas envisageable.

4.3 Détection des protéines de l'extrait de sol

Les analyses en spectrométrie de masse à partir de l'extrait protéique de sol ont permis d'identifier des protéines de *Pseudomonas* impliquées dans le fonctionnement général des cellules, et des enzymes du métabolisme général. Ce sont des protéines abondantes et largement répandues dans les cellules bactériennes. Les quelques protéines du sol identifiées appartiennent à la famille des chaperonnes.

Ces résultats peuvent s'expliquer de deux manières :

- (1) les seuils de détection actuels des spectromètres de masse nécessitent que les protéines soient issues d'environ 10^4 cellules d'une même espèce pour être détectées (von Bergen *et al.*, 2010). Les rendements d'extraction discutés au paragraphe 4.1 sont donc confrontés aux limites de la technique d'analyse.
- (2) la plupart des protéines ne sont pas identifiées parce qu'elles n'ont pas d'homologues dans les bases données. En effet, pour une protéine donnée, si les peptides détectés ne présentent pas 100% d'identité avec des séquences des bases, les peptides ne sont pas reconnus et la protéine correspondante n'est pas identifiable.

L'association de l'approche SIP et de la (meta)protéomique dans l'optique d'étudier la structure (phylogénétique), la fonction (potentiel métabolique) et l'activité (induction des voies métaboliques) des communautés bactériennes est une voie de recherche prometteuse. Cette approche requiert l'amélioration des techniques d'extraction de protéines du sol, le développement des performances des spectromètres de masse et de logiciels d'analyse adaptés, et surtout l'augmentation des données contenues dans les bases, par le biais du séquençage métagénomique des communautés bactériennes des sols.

V. Conclusions générales et perspectives du travail

Le principal objectif de ce projet était d'explorer la diversité des bactéries capables de dégrader les HAP dans un sol pollué de manière chronique par des hydrocarbures (bassin de rétention des eaux de chaussée d'autoroute). Notre démarche s'est résolument démarquée des approches classiques basées sur l'isolement de souches pures sur milieux sélectifs, dont on connaît les limites puisque moins de 5% des espèces bactériennes sont cultivables.

Nous avons choisi de mettre en œuvre une stratégie indépendante de la culture qui s'appuie sur le marquage métabolique par un isotope stable (^{13}C) des communautés bactériennes aptes à dégrader les HAP (méthode SIP). Associé à l'emploi d'outils moléculaires d'identification et à des analyses protéomiques, la méthode SIP présente l'avantage de coupler les informations phylogénétiques sur les communautés bactériennes à des données fonctionnelles relatives au métabolisme des HAP.

Dans cette optique, un HAP composé de trois cycles aromatiques, le phénanthrène, a été synthétisé sous forme uniformément marquée par le ^{13}C (Chapitre I). Cette molécule traceur a servi de substrat à des échantillons du sol incubés en microcosmes, et ce à deux concentrations différentes, 33 et 337 ppm (Chapitre II).

L'exploration de la diversité bactérienne par le séquençage de plusieurs centaines de gènes d'ARNr 16S, amplifiés à partir de l'ADN métagénomique marqué au ^{13}C , a permis d'identifier les populations spécifiquement impliquées dans la dégradation du phénanthrène. Nous avons ainsi montré que les Betaprotéobactéries étaient les principaux acteurs de cette dégradation, les taxons les plus abondants étant apparentés aux genres *Acidovorax*, *Rhodoferax*, *Polaromonas* et *Hydrogenophaga*. Quelques espèces voisines ont été isolées récemment par d'autres équipes et reconnues comme des acteurs importants de la dégradation des HAP. Nous avons aussi mis en évidence l'implication d'espèces de la famille des Rhodocyclaceae et du genre *Thiobacillus*, des taxons qui ont rarement été associés à la dégradation des HAP.

Ces résultats, confortés par des mesures de PCR quantitative (qPCR), soulignent le rôle majeur des Betaprotéobactéries dans la biodégradation des HAP *in situ*, et se démarquent des études antérieures sur des bactéries isolées, appartenant pour la plupart aux Sphingomonades (Alphaprotéobactéries), Pseudomonades (Gammaprotéobactéries) et Mycobactéries (Actinobactéries). Dans notre étude, les Alpha- et les Gammaprotéobactéries, ainsi que les Actinobactéries apparaissent comme des acteurs secondaires nettement moins abondants, sur

la base du pourcentage de séquences ARNr 16S et d'analyses par qPCR (Chapitre III). Les Gammaprotéobactéries que nous avons détectées, ne sont pas des *Pseudomonas* mais appartiennent à des taxons quasi inconnus de la classe des Xanthomonadales et à un taxon non affilié.

Nous avons aussi montré que sous l'effet du phénanthrène à 337 ppm, la structure des communautés bactériennes des microcosmes change, comme en témoignent les trois observations suivantes : (1) les populations capables de métaboliser les HAP se développent rapidement et le dégradent, (2) les populations sensibles aux effets toxiques du HAP sont inhibées, (3) lorsque le phénanthrène est dégradé, les communautés bactériennes tendent vers un retour à l'état initial (résilience). On observe une chute des effectifs des bactéries dégradant les HAP, qui pourrait s'expliquer par la prédation, et un retour à la normale du métabolisme global.

En outre, la concentration de phénanthrène ajouté aux microcosmes a un impact sur le temps de réponse des populations actives, mais pas sur la diversité des acteurs de la dégradation eux-mêmes.

Le deuxième volet de ce projet visait, par une approche moléculaire, à sonder la diversité phylogénétique et catalytique d'arène-dioxygénases, enzymes responsables de l'oxydation initiale des HAP.

Grâce à deux jeux d'amorces que nous avons conçus, plus d'une centaine de gènes de dioxygénases ont été amplifiés à partir de l'ADN marqué du sol. L'analyse phylogénétique des séquences a montré que les Sphingomonades étaient principalement impliquées dans la dégradation du phénanthrène, aux côtés des Burkholderiales. Les séquences détectées correspondent effectivement à des dioxygénases oxydant les HAP, mais celles-ci sont relativement distantes de celles des dioxygénases connues jusqu'à présent. La diversité observée est dépendante de la spécificité des amorces, et ne reflète que partiellement la richesse du sol. Plusieurs jeux d'amorces seraient nécessaires pour sonder cette diversité de façon plus exhaustive.

Nous avons ensuite développé un système d'expression visant à associer une activité catalytique aux séquences amplifiées de l'ADN du sol. Dans ce but, les dioxygénases de souches connues (*Sphingomonas* sp. CHY-1 et *Acidovorax* sp. NA3) ont servi de modèles pour la construction de dioxygénases hybrides. Cette stratégie a fait la preuve qu'il était

possible de connaître l'activité de dioxygénases du sol, en clonant seulement une séquence correspondant à son domaine catalytique. En revanche, lorsque cette séquence était trop divergente de celle de la dioxygénase modèle (moins de 80% de similarité), les protéines hybrides se sont révélées trop instables pour pouvoir doser leur activité.

Dans le but de construire d'autres hybrides avec la dioxygénase de *Acidovorax*. sp NA3 comme modèle, cette enzyme a été surproduite avec succès dans *E. coli* en lui associant les transporteurs d'électrons de *Sphingomonas* CHY-1.

Enfin, nous avons tenté de développer une approche protéomique dans le but de détecter certaines enzymes de dégradation des HAP directement à partir d'extraits du sol, notamment des dioxygénases. Un protocole d'extraction de protéines bactériennes du sol a été mis au point permettant un couplage avec l'analyse protéomique. Cette approche prometteuse requiert toutefois d'autres améliorations techniques avant de pouvoir analyser toute la richesse du metaprotéome bactérien par spectrométrie de masse.

Au cours de cette thèse, nous avons bien mis en évidence la grande diversité bactérienne des sols, et montré que les bactéries dégradant les HAP dans le sol étaient très mal connues. En effet, une proportion significative de séquences que nous avons obtenues (gène d'ARNr 16S et dioxygénases) étaient relativement distantes des taxons ou des familles d'enzymes connus.

Une des perspectives de ce travail serait d'entreprendre le séquençage métagénomique de l'ADN ^{13}C pour obtenir des informations sur le génome des espèces majoritaires impliquées dans la dégradation des HAP.

En employant les techniques de séquençage de nouvelle génération, on pourrait générer un volume de séquences suffisant pour reconstituer une partie du génome des acteurs majoritaires de la dégradation du phénanthrène. Ces données serviraient de base à des études approfondies des gènes cataboliques spécifiques du métabolisme des HAP, sans même avoir isolé les bactéries auxquelles ils appartiennent. En outre, ces résultats constitueraient une base de données précieuse pour l'identification, par l'approche protéomique, d'enzymes ou de protéines de fonction encore mal connue.

Par ailleurs, l'exploration de la diversité bactérienne que nous avons mise en œuvre a quasi exclusivement ciblé des microorganismes capables de dégrader des HAP facilement bioaccessibles. C'est peut-être une des raisons pour lesquelles les Actinobactéries, capables

pour dégrader les HAP adsorbés sur les particules de sol, représentent une part marginale de la diversité observée. Afin de mimer l'adsorption des HAP sur la matrice du sol, le phénanthrène marqué pourrait être déposé sur membrane hydrophobe avant d'être introduit dans le sol. Les bactéries capables de former des biofilms et de dégrader les HAP pourraient se développer sur ces membranes. Une analyse phylogénétique telle que celle qui a été mise en œuvre au cours de cette thèse, permettrait d'identifier d'autres populations bactériennes impliquées dans le métabolisme des HAP moins biodisponibles.

Enfin, l'approche développée pour étudier, sur le plan fonctionnel, les dioxygénases du sol nécessiterait encore des améliorations techniques afin d'obtenir un système robuste d'expression produisant des enzymes solubles et fonctionnelles. L'étude de plusieurs enzymes hybrides pourrait mettre en évidence le rôle de résidus impliqués dans le mécanisme catalytique, et surtout dans la spécificité vis à vis des HAP. A terme, on pourrait employer ce système pour faire une analyse fonctionnelle rapide des gènes de dioxygénases présents dans un échantillon environnemental.

Sur la base des séquences de gènes de dioxygénases obtenues, il serait intéressant de développer une méthode de PCR quantitative sur les ARN messagers marqués ^{13}C afin de connaître l'expression du potentiel catalytique des bactéries dégradant les HAP aux cours du temps, et en fonction de différents facteurs environnementaux : concentrations en phénanthrène, disponibilité en nutriments, température, salinité du milieu.

D'un point de vue plus appliqué, les résultats que nous avons obtenus pourraient permettre de développer des outils de diagnostic et de suivi des sols contaminés par les HAP. Connaissant les séquences d'ARNr 16S des bactéries dégradant les HAP, on peut envisager de les détecter *in situ* par une approche FISH (hybridation *in situ* par une sonde fluorescente) ou par qPCR en fonction des conditions environnementales : changements de température au cours des saisons, salinité des eaux en hiver, apports en oxygène et en nutriments par les roseaux. Appliqué à la gestion des bassins de retentions des eaux de chaussée, un tel suivi permettrait de mieux comprendre le système de traitement des eaux contaminées et d'améliorer l'efficacité du traitement par bioremédiation.

VI. Références Bibliographiques

- Ahmadian A, Ehn M, Hober S (2006). Pyrosequencing: History, biochemistry and future. *Clin. Chim. Acta* **363**: 83-94.
- Amann RI, Ludwig W, Schleifer KH (1995). Phylogenetic identification and *in-situ* detection of individual microbial-cells without cultivation. *Microbiol. Rev.* **59**: 143-169.
- Annweiler E, Materna A, Safinowski M, Kappler A, Richnow HH, Michaelis W *et al* (2000). Anaerobic degradation of 2-methylnaphthalene by a sulfate-reducing enrichment culture. *Appl. Environ. Microbiol.* **66**: 5329-5333.
- Ausubel FM, Brent R, Kingston R, Moore DD, Seidman JG, Smith JA *et al* (1999). *Short protocols in molecular biology*. Academic press.
- Baker GC, Smith JJ, Cowan DA (2003). Review and re-analysis of domain-specific 16S primers. *J. Microbiol. Method.* **55**: 541-555.
- Bakermans C, Madsen EL (2002). Diversity of 16S rDNA and naphthalene dioxygenase genes from coal-tar-waste-contaminated aquifer waters. *Microbial Ecol.* **44**: 95-106.
- Bakken LR (1985). Separation and purification of bacteria from soil. *Appl. Environ. Microbiol.* **49**: 1482-1487.
- Bastiaens L, Springael D, Wattiau P, Harms H, deWachter R, Verachtert H *et al* (2000). Isolation of adherent polycyclic aromatic hydrocarbon (PAH)-degrading bacteria using PAH-sorbing carriers. *Appl. Environ. Microbiol.* **66**: 1834-1843.
- Benndorf D, Balcke GU, Harms H, von Bergen M (2007). Functional metaproteome analysis of protein extracts from contaminated soil and groundwater. *ISME J.* **1**: 224-234.
- Bodour AA, Wang J-M, Brusseau ML, Maier RM (2003). Temporal change in culturable phenanthrene degraders in response to long-term exposure to phenanthrene in a soil column system. *Environ Microbiol* **5**: 888-95.
- Bombach P, Richnow HH, Kastner M, Fischer A (2010). Current approaches for the assessment of in situ biodegradation. *Appl. Microbiol. Biotechnol.* **86**: 839-852.
- Boonchan S, Britz ML, Stanley GA (2000). Degradation and mineralization of high-molecular-weight polycyclic aromatic hydrocarbons by defined fungal-bacterial cocultures. *Appl. Environ. Microbiol.* **66**: 1007-1019.
- Borde X, Guieysse B, Delgado O, Munoz R, Hatti-Kaul R, Nugier-Chauvin C *et al* (2003). Synergistic relationships in algal-bacterial microcosms for the treatment of aromatic pollutants. *Bioresource Technol.* **86**: 293-300.
- Bordenave S, Goni-urriza M, Vilette C, Blanchard S, Caumette P, Duran R (2008). Diversity of ring-hydroxylating dioxygenases in pristine and oil contaminated microbial mats at genomic and transcriptomic levels. *Environ Microbiol* **10**: 3201-3211.

- Bordenave S, Goni-Urriza MS, Caumette P, Duran R (2007). Effects of heavy fuel oil on the bacterial community structure of a pristine microbial mat. *Appl. Environ. Microbiol.* **73**: 6089-6097.
- Bosch R, Garcia-Valdes E, Moore ER (1999). Genetic characterization and evolutionary implications of a chromosomally encoded naphthalene-degradation upper pathway from *Pseudomonas stutzeri* AN10. *Gene* **236**: 149-57.
- Boschker HTS, Nold SC, Wellsbury P, Bos D, de Graaf W, Pel R *et al* (1998). Direct linking of microbial populations to specific biogeochemical processes by C-13-labelling of biomarkers. *Nature* **392**: 801-805.
- Bruley C, Dupierris V, Masselon C, Court M, Kieffer-Jaquinod S (2009). A toolbox for validation of mass spectrometry peptides identification and generation of database: IRMa. *Bioinformatics* **25**: 1980-1981.
- Carredano E, Karlsson A, Kauppi B, Choudhury D, Parales RE, Parales JV *et al* (2000). Substrate binding site of naphthalene 1,2-dioxygenase: functional implications of indole binding. *J. Mol. Biol.* **296**: 701-712.
- Cebron A, Beguiristain T, Faure P, Norini M-P, Masfaraud J-F, Leyval C (2009). Influence of vegetation on the *in situ* bacterial community and polycyclic aromatic hydrocarbon (PAH) degraders in aged PAH-contaminated or thermal-desorption-treated soil. *Appl. Environ. Microbiol.* **75**: 6322-6330.
- Cebron A, Norini MP, Beguiristain T, Leyval C (2008). Real-Time PCR quantification of PAH-ring hydroxylating dioxygenase (PAH-RHD alpha) genes from Gram positive and Gram negative bacteria in soil and sediment samples. *J. Microbiol. Meth.* **73**: 148-159.
- Cerniglia CE (1992). Biodegradation of polycyclic aromatic hydrocarbons. *Biodegradation* **3**: 351-368.
- Cerniglia CE (1997). Fungal metabolism of polycyclic aromatic hydrocarbons: Past, present and future applications in bioremediation. *J. Ind. Microbiol. Biotechnol.* **19**: 324-333.
- Cerniglia CE, Gibson DT, Vanbaalen C (1980). Oxidation of naphthalene by *Cyanobacteria* and microalgae. *J. Gen. Microbiol.* **116**: 495-500.
- Chadhain SMN, Moritz EM, Kim E, Zylstra GJ (2007). Identification, cloning, and characterization of a multicomponent biphenyl dioxygenase from *Sphingobium yanoikuyae* B1. *J. Ind. Microbiol. Biotechnol.* **34**: 605-613.
- Champreda V, Kanokratana P, Uengwetwanit T, Rattanachomsri U, Bunternngsook B, Nimchua T *et al* (2011). Insights into the phylogeny and metabolic potential of a primary tropical peat swamp forest microbial community by metagenomic analysis. *Microbial Ecol.* **61**: 518-528.
- Coates JD, Woodward J, Allen J, Philp P, Lovley DR (1997). Anaerobic degradation of polycyclic aromatic hydrocarbons and alkanes in petroleum-contaminated marine harbor sediments. *Appl. Environ. Microbiol.* **63**: 3589-3593.

- Cole JR, Konstantinidis K, Farris RJ, Tiedje JM (2010). Microbial diversity and phylogeny: extending from rRNAs to genomes. In: Liu WT and Jansson JK (eds). *Environmental Molecular Microbiology*. Caister Academic Press. pp 1-19.
- Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ *et al* (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic. Acids Res.* **37**: D141-D145.
- Correll CC, Batie CJ, Ballou DP, Ludwig ML (1992). Phthalate dioxygenase reductase: a modular structure for electron transfer from pyridine nucleotides to [2Fe-2S]. *Science* **258**: 1604-1610.
- Courtois S, Frostegard A, Goransson P, Depret G, Jeannin P, Simonet P (2001). Quantification of bacterial subgroups in soil: comparison of DNA extracted directly from soil or from cells previously released by density gradient centrifugation. *Environ. Microbiol.* **3**: 431-439.
- Criquet S, Farnet AM, Ferre E (2002). Protein measurement in forest litter. *Biol. Fertil. Soil* **35**: 307-313.
- Curtis TP, Sloan WT, Scannell JW (2002). Estimating prokaryotic diversity and its limits. *Proc Natl Acad Sci U S A* **99**: 10494-10499.
- David B. (1992). *Vol. PhD*. Université Blaise Pascal: Clermont-Ferrand II.
- De Lorenzo VL, Eltis B, Kessler, Timmis KN (1993). Analysis of *Pseudomonas* gene products using lacIq/P_{trp}-lac plasmids and transposons that confer conditional phenotypes. *Gene* **123**: 17-24.
- Delmont TO, Robe P, Cecillon S, Clark IM, Constancias F, Simonet P *et al* (2011). Accessing the soil metagenome for studies of microbial diversity. *Appl. Environ. Microbiol.* **77**: 1315-24.
- Demaneche S, Meyer C, Micoud J, Louwagie M, Willison JC, Jouanneau Y (2004). Identification and functional analysis of two aromatic-ring-hydroxylating dioxygenases from a *Sphingomonas* strain that degrades various polycyclic aromatic hydrocarbons. *Appl. Environ. Microbiol.* **70**: 6714-6725.
- Denome SA, Stanley DC, Olson ES, Young KD (1993). Metabolism of dibenzothiophene and naphthalene in *Pseudomonas* strains: complete DNA sequence of an upper naphthalene catabolic pathway. *J. Bacteriol.* **175**: 6890-901.
- DeRito CM, Pumphrey GM, Madsen EL (2005). Use of field-based stable isotope probing to identify adapted populations and track carbon flow through a phenol-degrading soil microbial community. *Appl. Environ. Microbiol.* **71**: 7858-65.
- Di Gregorio S, Zocca C, Sidler S, Toffanin A, Lizzari D, Vallini G (2004). Identification of two new sets of genes for dibenzothiophene transformation in *Burkholderia* sp. DBT1. *Biodegradation* **15**: 111-123.

- Diaz E (2004). Bacterial degradation of aromatic pollutants: a paradigm of metabolic versatility. *Int. Microbiol.* **7**: 173-180.
- Ding GC, Smalla K, Heuer H, Zuhlke S, Spiteller M, Pronk GJ *et al* (2010). Soil type-dependent responses to phenanthrene as revealed by determining the diversity and abundance of polycyclic aromatic hydrocarbon ring-hydroxylating dioxygenase genes by using a novel PCR detection system. *Appl. Environ. Microbiol.* **76**: 4765-4771.
- Doyle E, Muckian L, Hickey AM, Clipson N (2008). Microbial PAH degradation. *Adv. Appl. Microbiol.* **65**: 27-66.
- El Azhari N, Bru D, Sarr A, Martin-Laurent F (2008). Estimation of the density of the protocatechuate-degrading bacterial community in soil by real-time PCR. *Eur. J. Soil Sci.* **59**: 665-673.
- Epstein SS, Bollmann A, Lewis K (2007). Incubation of environmental samples in a diffusion chamber increases the diversity of recovered isolates. *Appl. Environ. Microbiol.* **73**: 6386-6390.
- Ferraro DJ, Gakhar L, Ramaswamy S (2005). Rieske business: Structure-function of Rieske non-heme oxygenases. *Biochem. Biophys. Res. Commun.* **338**: 175-190.
- Fierer N, Bradford MA, Jackson RB (2007). Toward an ecological classification of soil bacteria. *Ecology* **88**: 1354-1364.
- Fierer N, Jackson JA, Vilgalys R, Jackson RB (2005). Assessment of soil microbial community structure by use of taxon-specific quantitative PCR assays. *Appl Environ Microbiol* **71**: 4117-4120.
- Franklin FCH, Bagdasarian M, Bagdasarian MM, Timmis KN (1981). Molecular and functional-analysis of the tol plasmid pWWO from *Pseudomonas putida* and cloning of genes for the entire regulated aromatic ring meta-cleavage pathway. *Proc Natl Acad Sci USA* **78**: 7458-7462.
- Frostegard A, Courtois S, Ramisse V, Clerc S, Bernillon D, Le Gall F *et al* (1999). Quantification of bias related to the extraction of DNA directly from soils. *Appl. Environ. Microbiol.* **65**: 5409-5420.
- Fuenmayor SL, Wild M, Boyes AL, Williams PA (1998). A gene cluster encoding steps in conversion of naphthalene to gentisate in *Pseudomonas* sp. strain U2. *J. Bacteriol.* **180**: 2522-2530.
- Gandolfi I, Bestetti G, Siculo M, Franzetti A, Fontanarosa E, Santagostino A (2010). Influence of compost amendment on microbial community and ecotoxicity of hydrocarbon-contaminated soils. *Bioresource Technol.* **101**: 568-575.
- Gans J, Wolinsky M, Dunbar J (2005). Computational improvements reveal great bacterial diversity and high metal toxicity in soil. *Science* **309**: 1387-1390.

- Gascuel O, Dereeper A, Guignon V, Blanc G, Audic S, Buffet S *et al* (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucl. Acids Res.* **36**: W465-W469.
- Geiselbrecht AD, Herwig RP, Deming JW, Staley JT (1996). Enumeration and phylogenetic analysis of polycyclic aromatic hydrocarbon-degrading marine bacteria from Puget Sound sediments. *Appl. Environ. Microbiol.* **62**: 3344-3349.
- George KW, Hay AG (2011). Bacterial Strategies for Growth on Aromatic Compounds. *Adv. Appl. Microbiol.* **74**: 1-33.
- Gibson J, Harwood CS (2002). Metabolic diversity in aromatic compound utilization by anaerobic microbes. *Annu. Rev. Biochem.* **56**: 345-369.
- Gonzalez-y-Merchand JA, EstradaGarcia I, Colston MJ, Cox RA (1996). A novel method for the isolation of mycobacterial DNA. *FEMS Microbiol. Lett.* **135**: 71-77.
- Grimm AC, Harwood CS (1999). NahY, a catabolic plasmid-encoded receptor required for chemotaxis of *Pseudomonas putida* to the aromatic hydrocarbon naphthalene. *J. Bacteriol.* **181**: 3310-3316.
- Grosser RJ, Warshawsky D, Vestal JR (1991). Indigenous and enhanced mineralization of pyrene, benzo[a]pyrene, and carbazole in soils. *Appl. Environ. Microbiol.* **57**: 3462-3469.
- Grund AD, Gunsalus IC (1983). Cloning of genes for naphthalene metabolism in *Pseudomonas putida*. *J. Bacteriol.* **156**: 89-94.
- Habe H, Omori T (2003). Genetics of polycyclic aromatic hydrocarbon metabolism in diverse aerobic bacteria. *Biosci Biotech Biochem* **67**: 225-243.
- Hamann C, Hegemann J, Hildebrandt A (1999). Detection of polycyclic aromatic hydrocarbon degradation genes in different soil bacteria by polymerase chain reaction and DNA hybridization. *FEMS Microbiol. Lett.* **173**: 255-263.
- Han KD, Jung YT, Son SY (2003). Phylogenetic analysis of phenanthrene-degrading *Sphingomonas*. *J. Microbiol. Biotechnol.* **13**: 942-948.
- Haritash AK, Kaushik CP (2009). Biodegradation aspects of Polycyclic Aromatic Hydrocarbons (PAHs): a review. *J. Hazard. Mater.* **169**: 1-15.
- Hazen TC, Dubinsky EA, DeSantis TZ, Andersen GL, Piceno YM, Singh N *et al* (2010). Deep-sea oil plume enriches indigenous oil-degrading bacteria. *Science* **330**: 204-208.
- Heitkamp MA, Freeman JP, Miller DW, Cerniglia CE (1988). Pyrene degradation by a *Mycobacterium* sp.: identification of ring oxidation and ring fission products. *Appl. Environ. Microbiol.* **54**: 2556-2565.
- Hill GT, Mitkowski NA, Aldrich-Wolfe L, Emele LR, Jurkonie DD, Ficke A *et al* (2000). Methods for assessing the composition and diversity of soil microbial communities. *Appl. Soil Ecol.* **15**: 25-36.

- Hiraishi A, Yonemitsu Y, Matsushita M, Shin YK, Kuraishi H, Kawahara K (2002). Characterization of *Porphyrobacter sanguineus* sp nov., an aerobic bacteriochlorophyll-containing bacterium capable of degrading biphenyl and dibenzofuran. *Arch. Microbiol.* **178**: 45-52.
- Hosse M, Wilkinson KJ (2001). Determination of electrophoretic mobilities and hydrodynamic radii of three humic substances as a function of pH and ionic strength. *Environ. Sci. Technol.* **35**: 4301-4306.
- Hugenholtz P (2002). Exploring prokaryotic diversity in the genomic era. *Genome Biol.* **3**: reviews0003-reviews0003.8.
- Hugenholtz P, Tyson GW (2008). Microbiology - Metagenomics. *Nature* **455**: 481-483.
- Iwabuchi T, Harayama S (1998). Biochemical and molecular characterization of 1-hydroxy-2-naphthoate dioxygenase from *Nocardioides* sp. KP7. *J. Biol. Chem.* **273**: 8332-8336.
- Iwai S, Johnson TA, Chai BL, Hashsham SA, Tiedje JM (2011). Comparison of the specificities and efficacies of primers for aromatic dioxygenase gene analysis of environmental samples. *Appl. Environ. Microbiol.* **77**: 3551-3557.
- Jain RK, Chauhan A (2010). Biodegradation: gaining insight through proteomics. *Biodegradation* **21**: 861-879.
- Jakoncic J, Jouanneau Y, Meyer C, Stojanoff V (2007a). The catalytic pocket of the ring-hydroxylating dioxygenase from *Sphingomonas* CHY-1. *Biochem. Biophys. Res. Commun.* **352**: 861-866.
- Jakoncic J, Jouanneau Y, Meyer C, Stojanoff V (2007b). The crystal structure of the ring-hydroxylating dioxygenase from *Sphingomonas* CHY-1. *FEBS J* **274**: 2470-2481.
- Janssen PH (2006). Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Appl. Environ. Microbiol.* **72**: 1719-1728.
- Janssen PH, Yates PS, Grinton BE, Taylor PM, Sait M (2002). Improved culturability of soil bacteria and isolation in pure culture of novel members of the divisions Acidobacteria, Actinobacteria, Proteobacteria, and Verrucomicrobia. *Appl. Environ. Microbiol.* **68**: 2391-2396.
- Jansson JK, Hjort K, Lembke A, Speksnijder A, Smalla K (2007). Community structure of actively growing bacterial populations in plant pathogen suppressive soil. *Microbial Ecol.* **53**: 399-413.
- Jehmlich N, Schmidt F, Taubert M, Seifert J, von Bergen M, Richnow HH *et al* (2009). Comparison of methods for simultaneous identification of bacterial species and determination of metabolic activity by protein-based stable isotope probing (Protein-SIP) experiments. *Rapid Commun. Mass Spectrom.* **23**: 1871-1878.

- Jehmlich N, Schmidt F, von Bergen M, Richnow HH, Vogt C (2008). Protein-based stable isotope probing (Protein-SIP) reveals active species within anoxic mixed cultures. *ISME J.* **2**: 1122-1133.
- Jeon CO, Park M, Ro HS, Park W, Madsen EL (2006). The naphthalene catabolic (nag) genes of *Polaromonas naphthalenivorans* CJ2: Evolutionary implications for two gene clusters and novel regulatory control. *Appl Environ Microbiol* **72**: 1086-1095.
- Jeon CO, Park W, Padmanabhan P, DeRito CJ, Snape R, Madsen EL (2003). Discovery of a bacterium, with distinctive dioxygenase, that is responsible for in situ biodegradation in contaminated sediment. *Proc Natl Acad Sci USA* **100** 13591-13596.
- Johnsen AR, Wick LY, Harms H (2005). Principles of microbial PAH-degradation in soil. *Environ. Pollut.* **133**: 71-84.
- Johnsen AR, Winding A, Karlson U, Roslev P (2002). Linking of microorganisms to phenanthrene metabolism in soil by analysis of C-13-labeled cell lipids. *Appl. Environ. Microbiol.* **68**: 6106-6113.
- Jones MD, Singleton DR, Sun W, Aitken MD (2011). Multiple DNA extractions coupled with stable-isotope probing of anthracene-degrading bacteria in contaminated soil. *Appl. Environ. Microbiol.* **77**: 2984-2991.
- Jouanneau Y (2010). Oxidative inactivation of ring-cleavage extradiol dioxygenases: mechanism and ferredoxin-mediated reactivation. In: Timmis KN (ed). *Microbiology of hydrocarbons, oils, lipids, and derived compounds*. Springer: New York. pp 1–10.
- Jouanneau Y, Martin F, Krivobok S, Willison JC (2011). Ring-hydroxylating dioxygenases involved in PAH degradation : structure, function and biodiversity. . In: Koukkou AI (ed). *In Microbial biodegradation of non metals: current approaches*. Horizon Scientific Press.
- Jouanneau Y, Meyer C (2006). Purification and characterization of an arene cis-dihydrodiol dehydrogenase endowed with broad substrate specificity toward polycyclic aromatic hydrocarbon dihydrodiols. *Appl. Environ. Microbiol.* **72**: 4726-4734.
- Jouanneau Y, Meyer C, Jakoncic J, Stojanoff V, Gaillard J (2006). Characterization of a naphthalene dioxygenase endowed with an exceptionally broad substrate specificity toward polycyclic aromatic hydrocarbons. *Biochemistry* **45**: 12380-12391.
- Jouanneau Y, Micoud J, Meyer C (2007). Purification and characterization of a three-component salicylate 1-hydroxylase from *Sphingomonas* sp. strain CHY-1. *Appl. Environ. Microbiol.* **73**: 7515-7521.
- Jouanneau Y, Willison JC, Meyer C, Krivobok S, Chevron N, Besombes J *et al* (2005). Stimulation of pyrene mineralization in freshwater sediments by bacterial and plant bioaugmentation. *Environ. Sci. Technol.* **39**: 5229-5235.
- Jouanneau Y, Willison JC, Rodarie D (1999). Dégradation microbiologique des hydrocarbure aromatiques polycycliques. *Rapport pour l'Agence de l'Environnement et de la Maitrise de l'Energie, CEA Grenoble*.

- Juhas M, van der Meer JR, Gaillard M, Harding RM, Hood DW, Crook DW (2009). Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol. Rev.* **33**: 376-393.
- Juhasz AL, Stanley GA, Britz ML (2000). Microbial degradation and detoxification of high molecular weight polycyclic aromatic hydrocarbons by *Stenotrophomonas maltophilia* strain VUN 10,003. *Lett. Appl. Microbiol.* **30**: 396-401.
- Kahl S, Hofer B (2003). A genetic system for the rapid isolation of aromatic-ring-hydroxylating dioxygenase activities. *Microbiology* **149**: 1475-1481.
- Karlsson A, Parales JV, Parales RE, Gibson DT, Eklund H, Ramaswamy S (2003). Crystal structure of naphthalene dioxygenase: side-on binding of dioxygen to iron. *Science* **299**: 1039-1042.
- Kasai Y, Shindo K, Harayama S, Misawa N (2003). Molecular characterization and substrate preference of a polycyclic aromatic hydrocarbon dioxygenase from *Cycloclasticus* sp strain A5. *Appl Environ Microbiol* **69**: 6688-6697.
- Kauppi B, Lee K, Carredano E, Parales RE, Gibson DT, Eklund H *et al* (1998). Structure of an aromatic-ring-hydroxylating dioxygenase-naphthalene 1,2- dioxygenase. *Structure* **6**: 571-586.
- Kelly DP, Wood AP, Stackebrandt E (2005). Genus II. *Thiobacillus*. In: Brenner DJ, Krieg NR and Staley JT (eds). *Bergey's Manual of Systematic Bacteriology - the Alpha-, Beta-, Delta-, and Epsilonproteobacteria*. Springer: New York. pp 764-769.
- Khan AA, Wang RF, Cao WW, Doerge DR, Wennerstrom D, Cerniglia CE (2001). Molecular cloning, nucleotide sequence, and expression of genes encoding a polycyclic aromatic ring dioxygenase from *Mycobacterium* sp strain PYR-1. *Appl. Environ. Microbiol.* **67**: 3577-3585.
- Kim SJ, Kweon O, Jones RC, Edmondson RD, Cerniglia CE (2008). Genomic analysis of polycyclic aromatic hydrocarbon degradation in *Mycobacterium vanbaalenii* PYR-1. *Biodegradation* **19**: 859-881.
- Kim SJ, Kweon O, Jones RC, Freeman JP, Edmondson RD, Cerniglia CE (2007). Complete and integrated pyrene degradation pathway in *Mycobacterium vanbaalenii* PYR-1 based on systems biology. *J. Bacteriol.* **189**: 464-472.
- Kim SJ, Kweon OG, Freeman JP, Jones RC, Adjei MD, Jhoo JW *et al* (2006). Molecular cloning and expression of genes encoding a novel dioxygenase Involved in low- and high-molecular-weight polycyclic aromatic hydrocarbon degradation in *Mycobacterium vanbaalenii* PYR-1. *Appl Environ Microbiol* **72**: 1045 - 1054.
- Kirk JL, Beaudette LA, Hart M, Moutoglis P, Khironomos JN, Lee H *et al* (2004). Methods of studying soil microbial diversity. *J. Microbiol. Method.* **58**: 169-188.
- Kiyohara H, Nagao K, Kouno K, Yano K (1982). Phenanthrene-degrading phenotype of *Alcaligenes faecalis* AFK2. *Appl Environ Microbiol* **43**: 458 - 461.

- Klock HE, Koesema EJ, Knuth MW, Lesley SA (2008). Combining the polymerase incomplete primer extension method for cloning and mutagenesis with microscreening to accelerate structural genomics efforts. *Proteins* **71**: 982-994.
- Kovaleva EG, Lipscomb JD (2008). Versatility of biological non-heme Fe(II) centers in oxygen activation reactions. *Nature Chem. Biol.* **4**: 186-193.
- Krivobok S, Kuony S, Meyer C, Louwagie M, Willison JC, Jouanneau Y (2003). Identification of pyrene-induced proteins in *Mycobacterium* sp. strain 6PY1: evidence for two ring-hydroxylating dioxygenases. *J. Bacteriol.* **185**: 3828-41.
- Kurkela S, Lehvaslaiho H, Palva ET, Teeri TH (1988). Cloning, nucleotide sequence and characterization of genes encoding naphthalene dioxygenase of *Pseudomonas putida* strain NCIB9816. *Gene* **73**: 355-62.
- Kweon O, Kim S-J, Baek S, Chae J-C, Adjei M, Baek D-H *et al* (2008). A new classification system for bacterial Rieske non-heme iron aromatic ring-hydroxylating oxygenases. *BMC Biochem.* **9**: 11.
- Kweon O, Kim SJ, Jones RC, Freeman JP, Adjei MD, Edmondson RD *et al* (2007). A polyomic approach to elucidate the fluoranthene-degradative pathway in *Mycobacterium vanbaalenii* PYR-1. *J. Bacteriol.* **189**: 4635-4647.
- Laemmli UK (1970). Cleavage of structural proteins during assembly of head of bacteriophage-T4. *Nature* **227**: 680-&.
- Laurie AD, Lloyd-Jones G (1999). The *phn* genes of *Burkholderia* sp. strain RP007 constitute a divergent gene cluster for polycyclic aromatic hydrocarbon catabolism. *J. Bacteriol.* **181**: 531-540.
- Law AMJ, Aitken MD (2003). Bacterial chemotaxis to naphthalene desorbing from a nonaqueous liquid. *Appl. Environ. Microbiol.* **69**: 5968-5973.
- Leys NM, Ryngaert A, Bastiaens L, Wattiau P, Top EM, Verstraete W *et al* (2005). Occurrence and community composition of fast-growing *Mycobacterium* in soils contaminated with polycyclic aromatic hydrocarbons. *FEMS Microbiol Ecol* **51**: 375-88.
- Leys NMEJ, Ryngaert A, Bastiaens L, Verstraete W, Top EM, Springael D (2004). Occurrence and phylogenetic diversity of *Sphingomonas* strains in soils contaminated with polycyclic aromatic hydrocarbons. *Appl Environ Microbiol* **70**: 1944-1955.
- Lillis L, Doyle E, Clipson N (2009). Comparison of DNA- and RNA-based bacterial community structures in soil exposed to 2,4-dichlorophenol. *J. Appl. Microbiol.* **107**: 1883-1893.
- Lindahl V (1996). Improved soil dispersion procedures for total bacterial counts, extraction of indigenous bacteria and cell survival. *J. Microbiol. Meth.* **25**: 279-286.

- Lopez-Gutierrez JC, Henry S, Hallet S, Martin-Laurent F, Catroux G, Philippot L (2004). Quantification of a novel group of nitrate-reducing bacteria in the environment by real-time PCR. *J. Microbiol. Meth.* **57**: 399-407.
- Ludwig W, Klenk H-P (2001). Overview: a phylogenetic backbone and taxonomic framework for procaryotic systematics. In: R.W. BDRaC (ed). *Bergey's manual of systematic bacteriology*, Springer-Verlag, New York, N.Y. edn. pp 49-65.
- Lueders T, Manefield M, Friedrich MW (2004). Enhanced sensitivity of DNA- and rRNA-based stable isotope probing by fractionation and quantitative analysis of isopycnic centrifugation gradients. *Environ Microbiol* **6**: 73-78.
- Magic-Knezev A, Wullings B, Van der Kooij D (2009). Polaromonas and Hydrogenophaga species are the predominant bacteria cultured from granular activated carbon filters in water treatment. *J. Appl. Microbiol.* **107**: 1457-1467.
- Mahmood S, Paton GI, Prosser JI (2005). Cultivation-independent in situ molecular analysis of bacteria involved in degradation of pentachlorophenol in soil. *Environ. Microbiol.* **7**: 1349-1360.
- Manefield M, Whiteley AS, Griffiths RI, Bailey MJ (2002). RNA stable isotope probing, a novel means of linking microbial community function to Phylogeny. *Appl Environ Microbiol* **68**: 5367-5373.
- Mannisto MK, Tirola MA, Salkinoja-Salonen MS, Kulomaa MS, Puhakka JA (1999). Diversity of chlorophenol-degrading bacteria isolated from contaminated boreal groundwater. *Arch. Microbiol.* **171**: 189-197.
- Maron PA, Mougél C, Siblot S, Abbas H, Lemanceau P, Ranjard L (2007). Protein extraction and fingerprinting optimization of bacterial communities in natural environment. *Microbial Ecol.* **53**: 426-434.
- Maron PA, Schimann H, Ranjard L, Brothier E, Domenach AM, Lensi R *et al* (2006). Evaluation of quantitative and qualitative recovery of bacterial communities from different soil types by density gradient centrifugation. *Eur. J. Soil Biol.* **42**: 65-73.
- Martin-Laurent F, Philippot L, Hallet S, Chaussod R, Germon JC, Soulas G *et al* (2001). DNA extraction from soils: Old bias for new microbial diversity analysis methods (vol 67, pg 2354, 2001). *Appl. Environ. Microbiol.* **67**: 4397-4397.
- Martineau C, Greer CW, Whyte LG (2010). Stable isotope probing analysis of the diversity and activity of methanotrophic bacteria in soils from the canadian high arctic. *Appl. Environ. Microbiol.* **76**: 5773-5784.
- Martineau C, Whyte LG, Greer CW (2008). Development of a SYBR safe (TM) technique for the sensitive detection of DNA in cesium chloride density gradients for stable isotope probing assays. *J Microbiol Methods* **73**: 199-202.
- Maughan H, Van der Auwera G (2011). Bacillus taxonomy in the genomic era finds phenotypes to be essential though often misleading. *Infect. Genet. Evol.* **11**: 789-797.

- Mayr C, Winding A, Hendriksen NB (1999). Community level physiological profile of soil bacteria unaffected by extraction method. *J. Microbiol. Method.* **36**: 29-33.
- Meckenstock RU, Annweiler E, Michaelis W, Richnow HH, Schink B (2000). Anaerobic naphthalene degradation by a sulfate-reducing enrichment culture. *Appl. Environ. Microbiol.* **66**: 2743-2747.
- Meckenstock RU, Safinowski M, Griebler C (2004). Anaerobic degradation of polycyclic aromatic hydrocarbons. *FEMS Microbiol Ecol* **49**: 27-36.
- Medvedev P, Stanciu M, Brudno M (2009). Computational methods for discovering structural variation with next-generation sequencing. *Nature Meth.* **6**: S13-20.
- Menn FM, Applegate BM, Sayler GS (1993). Nah plasmid-mediated catabolism of anthracene and phenanthrene to naphthoic acids. *Appl. Environ. Microbiol.* **59**: 1938-1942.
- Metzker ML (2010). Applications of next-generation sequencing. Sequencing technologies - the next generation. *Nat. Rev. Genet.* **11**: 31-46.
- Meulenbergh R, Rijnaarts HHM, Doddema HJ, Field JA (1997). Partially oxidized polycyclic aromatic hydrocarbons show an increased bioavailability and biodegradability. *FEMS Microbiol. Lett.* **152**: 45-49.
- Miyata N, Iwahori K, Foght JM, Gray MR (2004). Saturable, energy-dependent uptake of phenanthrene in aqueous phase by *Mycobacterium* sp strain RJGII-135. *Appl. Environ. Microbiol.* **70**: 363-369.
- Moody JD, Freeman JP, Doerge DR, Cerniglia CE (2001). Degradation of phenanthrene and anthracene by cell suspensions of *Mycobacterium* sp strain PYR-1. *Appl. Environ. Microbiol.* **67**: 1476-1483.
- Moreno-Ruiz E, Hernaez MJ, Martinez-Perez O, Santero E (2003). Identification and functional characterization of *Sphingomonas macrogolita* strain TFA genes involved in the first two steps of the tetralin catabolic pathway. *J Bacteriol* **185**: 2026-30.
- Moser R, Stahl U (2001). Insights into the genetic diversity of initial dioxygenases from PAH-degrading bacteria. *Appl. Microbiol. Biotechnol.* **55**: 609-618.
- Mouillot D, Lepretre A (1999). A comparison of species diversity estimators. *Res. Population Ecol.* **41**: 203-215.
- Muhling M, Woolven-Allen J, Murrell JC, Joint I (2008). Improved group-specific PCR primers for denaturing gradient gel electrophoresis analysis of the genetic diversity of complex microbial communities. *ISME J* **2**: 379-392.
- Nam JW, Nojiri H, Noguchi H, Uchimura H, Yoshida T, Habe H *et al* (2002). Purification and characterization of carbazole 1,9a-dioxygenase, a three-component dioxygenase system of *Pseudomonas resinovorans* strain CA10. *Appl Environ Microbiol* **68**: 5882-90.

- Nam JW, Nojiri H, Yoshida T, Habe H, Yamane H, Omori T (2001). New classification system for oxygenase components involved in ring- hydroxylating oxygenations. *Biosci. Biotechnol. Biochem.* **65**: 254-263.
- Narro ML, Cerniglia CE, Vanbaalen C, Gibson DT (1992). Metabolism of phenanthrene by the marine *Cyanobacterium agmenellum-quadruplicatum* Pr-6. *Appl. Environ. Microbiol.* **58**: 1351-1359.
- Nelson KE, Bryan PA, White BA (2010). Genomics and metagenomics: History and progress. In: Liu WT and Jansson JK (eds). *Environmental Molecular Microbiology*. Caister Academic Press. pp 21-36.
- Nemir A, David MM, Perrussel R, Sapkota A, Simonet P, Monier JM *et al* (2010). Comparative phylogenetic microarray analysis of microbial communities in TCE-contaminated soils. *Chemosphere* **80**: 600-607.
- Neufeld JD, Dumont MG, Vohra J, Murrell JC (2007). Methodological considerations for the use of stable isotope probing in microbial ecology. *Microbial Ecol.* **53**: 435-442.
- Neufeld JD, Mohn WW, de Lorenzo V (2006). Composition of microbial communities in hexachlorocyclohexane (HCH) contaminated soils from Spain revealed with a habitat-specific microarray. *Environ. Microbiol.* **8**: 126-140.
- Ono A, Miyazaki R, Sota M, Ohtsubo Y, Nagata Y, Tsuda M (2007). Isolation and characterization of naphthalene-catabolic genes and plasmids from oil-contaminated soil by using two cultivation-independent approaches. *Appl. Microbiol. Biotechnol.* **74**: 501-510.
- Paisse S, Coulon F, Goni-Urriza M, Peperzak L, McGenity TJ, Duran R (2008). Structure of bacterial communities along a hydrocarbon contamination gradient in a coastal sediment. *FEMS Microbiol. Ecol.* **66**: 295-305.
- Paisse S, Goni-Urriza M, Coulon F, Duran R (2010). How a bacterial community originating from a contaminated coastal sediment responds to an oil input. *Microb. Ecol.* **60**: 394-405.
- Parales RE, Lee K, Resnick SM, Jiang HY, Lessner DJ, Gibson DT (2000). Substrate specificity of naphthalene dioxygenase: Effect of specific amino acids at the active site of the enzyme. *J. Bacteriol.* **182**: 1641-1649.
- Parales RE, Resnick SM (2006). Aromatic ring hydroxylating dioxygenases. In: Ramos JL and Levesque RC (eds). *Pseudomonas*. Springer. pp 287-340.
- Peña-Méndez EM, Havel J, Patočka J (2005). Humic substances - compounds of still unknown structure: applications in agriculture, industry, environment, and biomedicine. *J. Appl. Biomed.* **3**: 13-24.
- Peng JJ, Cai C, Qiao M, Li H, Zhu YG (2010). Dynamic changes in functional gene copy numbers and microbial communities during degradation of pyrene in soils. *Environ. Pollut.* **158**: 2872-2879.

- Peng RH, Xiong AS, Xue Y, Fu XY, Gao F, Zhao W *et al* (2008). Microbial biodegradation of polyaromatic hydrocarbons. *FEMS Microbiol Rev* **32**: 927-55.
- Peyret P, Militon C, Boucher D, Vachelard C, Perchet G, Barra V *et al* (2010). Bacterial community changes during bioremediation of aliphatic hydrocarbon-contaminated soil. *FEMS Microbiol Ecol* **74**: 669-681.
- Philippot L, Andersson SGE, Battin TJ, Prosser JI, Schimel JP, Whitman WB *et al* (2010a). The ecological coherence of high bacterial taxonomic ranks. *Nat. Rev. Microbiol.* **8**: 523-529.
- Philippot L, Tscherko D, Bru D, Kandeler E (2010b). Distribution of high bacterial taxa across the chronosequence of two alpine glacier forelands. *Microb Ecol* **61**: 303-12.
- Pinyakong O, Habe H, Kouzuma A, Nojiri H, Yamane H, Omori T (2004). Isolation and characterization of genes encoding polycyclic aromatic hydrocarbon dioxygenase from acenaphthene and acenaphthylene degrading *Sphingomonas* sp strain A4. *FEMS Microbiol. Lett.* **238**: 297-305.
- Pinyakong O, Habe H, Omori T (2003a). The unique aromatic catabolic genes in sphingomonads degrading polycyclic aromatic hydrocarbons (PAHs). *J. Gen. Appl. Microbiol.* **49**: 1-19.
- Pinyakong O, Habe H, Yoshida T, Nojiri H, Omori T (2003b). Identification of three novel salicylate 1-hydroxylases involved in the phenanthrene degradation of *Sphingobium* sp. strain P2. *Biochem Biophys Res Commun* **301**: 350 - 357.
- Popp N, Schlomann M, Mau M (2006). Bacterial diversity in the active stage of a bioremediation system for mineral oil hydrocarbon-contaminated soils. *Microbiology* **152**: 3291-3304.
- Powell SN, Aitken MD, Singleton DR (2008). Effects of enrichment with salicylate on bacterial selection and PAH mineralization in a microbial community from a bioreactor treating contaminated soil. *Environ. Sci. Technol.* **42**: 4099-4105.
- Prosser J, Jansson JK, Liu W (2010). Nucleic-acid-based characterization of community structure and function. In: Liu WT and Jansson JK (eds). *Environmental Molecular Microbiology*. Caister Academic Press. pp 63-86.
- Pumphrey GM, Madsen EL (2008a). Field-based stable isotope probing reveals the identities of benzoic acid-metabolizing microorganisms and their in situ growth in agricultural soil. *Appl. Environ. Microbiol.* **74**: 4111-4118.
- Pumphrey GM, Madsen EL (2008b). Field-based stable isotope probing reveals the identities of benzoic acid-metabolizing microorganisms and their in situ growth in agricultural soil. *Appl Environ Microbiol* **74**: 4111-4118.
- Pumphrey GM, Ranchou-Peyruse A, Spain JC (2011). Cultivation-independent detection of autotrophic hydrogen-oxidizing bacteria by DNA stable-isotope probing. *Appl. Environ. Microbiol.* **77**: 4931-38.

- Rabus R, Widdel F (1995). Anaerobic degradation of ethylbenzene and other aromatic-hydrocarbons by new denitrifying bacteria. *Arch Microbiol* **163**: 96-103.
- Radajewski S, Ineson P, Parekh NR, Murrell JC (2000). Stable-isotope probing as a tool in microbial ecology. *Nature* **403**: 646-649.
- Radajewski S, McDonald IR, Murrell JC (2003). Stable-isotope probing of nucleic acids: a window to the function of uncultured microorganisms. *Curr. Opin. Biotechnol.* **14**: 296-302.
- Rafii F, Butler WR, Cerniglia CE (1992). Differentiation of a rapidly growing, scotochromogenic, polycyclic-aromatic-hydrocarbon-metabolizing strain of *Mycobacterium* sp from other known *Mycobacterium* species. *Arch. Microbiol.* **157**: 512-520.
- Rajendhran J, Gunasekaran P (2008). Strategies for accessing soil metagenome for desired applications. *Biotechnol. Adv.* **26**: 576-590.
- Ramette A (2007). Multivariate analyses in microbial ecology. *FEMS Microbiol Ecol* **62**: 142-160.
- Rappe MS, Giovannoni SJ (2003). The uncultured microbial majority. *Ann. Rev. Microbiol.* **57**: 369-394.
- Rentz JA, Alvarez PJJ, Schnoor JL (2005). Benzo[a]pyrene co-metabolism in the presence of plant root extracts and exudates: Implications for phytoremediation. *Environ. Pollut.* **136**: 477-484.
- Robe P, Nalin R, Capellano C, Vogel TA, Simonet P (2003). Extraction of DNA from soil. *Eur. J. Soil Biol.* **39**: 183-190.
- Rockne KJ, Chee-Sanford JC, Sanford RA, Hedlund BP, Staley JT, Strand SE (2000). Anaerobic naphthalene degradation by microbial pure cultures under nitrate-reducing conditions. *Appl. Environ. Microbiol.* **66**: 1595-1601.
- Rodriguez-Blanco A, Ghiglione JF, Catala P, Casamayor EO, Lebaron P (2009). Spatial comparison of total vs. active bacterial populations by coupling genetic fingerprinting and clone library analyses in the NW Mediterranean Sea. *FEMS Microbiol Ecol* **67**: 30-42.
- Rodriguez-Valera F, Martin-Cuadrado AB, Rodriguez-Brito B, Pasic L, Thingstad TF, Rohwer F *et al* (2009). Explaining microbial population genomics through phage predation. *Nature Rev. Microbiol.* **7**: 828-836.
- Roesch LF, Fulthorpe RR, Riva A, Casella G, Hadwin AKM, Kent AD *et al* (2007). Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME J.* **1**: 283-290.
- Romero MC, Cazau MC, Giorgieri S, Arambarri AM (1998). Phenanthrene degradation by microorganisms isolated from a contaminated stream. *Environ. Pollut.* **101**: 355-359.
- Romine MF, Stillwell LC, Wong KK, Thurston SJ, Sisk EC, Sensen C *et al* (1999). Complete sequence of a 184-kilobase catabolic plasmid from *Sphingomonas aromaticivorans* F199. *J Bacteriol* **181**: 1585 - 1602.

- Ronaghi M, Uhlen M, Nyren P (1998). A sequencing method based on real-time pyrophosphate. *Science* **281**: 363-+.
- Rotaru A-E, Probian C, Wilkes H, Harder J (2010). Highly enriched Betaproteobacteria growing anaerobically with p-xylene and nitrate. *FEMS Microbiol Ecol* **71**: 460-8.
- Roussel JC (1996). Analyse et Caractérisation *Technique de l'ingénieur* pp 1-23.
- Safinowski M, Meckenstock RU (2006). Methylation is the initial reaction in anaerobic naphthalene degradation by a sulfate-reducing enrichment culture. *Environ. Microbiol.* **8**: 347-352.
- Sait M, Hugenholtz P, Janssen PH (2002). Cultivation of globally distributed soil bacteria from phylogenetic lineages previously only detected in cultivation-independent surveys. *Environ Microbiol* **4**: 654-666.
- Saito A, Iwabuchi T, Harayama S (2000). A novel phenanthrene dioxygenase from *Nocardioides* sp strain KP7: Expression in *Escherichia coli*. *J. Bacteriol.* **182**: 2134-2141.
- Samanta SK, Singh OV, Jain RK (2002). Polycyclic aromatic hydrocarbons: environmental pollution and bioremediation. *Trends Biotechnol.* **20**: 243-248.
- Sanger F, Nicklen S, Coulson AR (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* **74**: 5463-7.
- Schloss PD, Handelsman J (2004). Status of the microbial census. *Microbiol. Mol. Biol. Rev.* **68**: 686-+.
- Schloss PD, Handelsman J (2005). Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl. Environ. Microbiol.* **71**: 1501-1506.
- Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB *et al* (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **75**: 7537-7541.
- Schneider J, Grosser R, Jayasimhulu K, Xue W, Warshawsky D (1996). Degradation of pyrene, benz[a]anthracene, and benzo[a]pyrene by *Mycobacterium* sp strain RJGII-135, isolated from a former coal gasification site. *Appl. Environ. Microbiol.* **62**: 13-19.
- Schuler L, Chadhain SMN, Jouanneau Y, Meyer C, Zylstra GJ, Hols P *et al* (2008). Characterization of a novel angular dioxygenase from fluorene-degrading *Spingomonas* sp strain LB126. *Appl. Environ. Microbiol.* **74**: 1050-1057.
- Schuler L, Jouanneau Y, Chadhain SMN, Meyer C, Pouli M, Zylstra GJ *et al* (2009). Characterization of a ring-hydroxylating dioxygenase from phenanthrene-degrading *Spingomonas* sp strain LH128 able to oxidize benz[a]anthracene. *Appl. Microbiol. Biotechnol.* **83**: 465-475.

- Senda T, Yamada T, Sakurai N, Kubota M, Nishizaki T, Masai E *et al* (2000). Crystal structure of NADH-dependent ferredoxin reductase component in biphenyl dioxygenase. *J. Mol. Biol.* **304**: 397-410.
- Seo JS, Keum YS, Hu YT, Lee SE, Li QX (2007). Degradation of phenanthrene by *Burkholderia* sp C3: initial 1,2- and 3,4-dioxygenation and meta- and ortho-cleavage of naphthalene-1,2-diol. *Biodegradation* **18**: 123-131.
- Shendure J, Ji HL (2008). Next-generation DNA sequencing. *Nat. Biotechnol.* **26**: 1135-1145.
- Shi T, Fredrickson JK, Balkwill DL (2001). Biodegradation of polycyclic aromatic hydrocarbons by *Sphingomonas* strains isolated from the terrestrial subsurface. *J. Ind. Microbiol. Biotechnol.* **26**: 283-289.
- Shuttleworth KL, Cerniglia CE (1996). Bacterial degradation of low concentrations of phenanthrene and inhibition by naphthalene. *Microbial Ecol.* **31**: 305-317.
- Simon MJ, Osslund TD, Saunders R, Ensley BD, Suggs S, Harcourt A *et al* (1993). Sequences of genes encoding naphthalene dioxygenase in *Pseudomonas putida* strains G7 and NCIB 9816-4. *Gene* **127**: 31-7.
- Singleton DR, Hunt M, Powell SN, Frontera-Suau R, Aitken MD (2007). Stable-isotope probing with multiple growth substrates to determine substrate specificity of uncultivated bacteria. *J. Microbiol. Method.* **69**: 180-187.
- Singleton DR, Powell SN, Sangaiah R, Gold A, Ball LM, Aitken MD (2005). Stable-isotope probing of bacteria capable of degrading salicylate, naphthalene, or phenanthrene in a Bioreactor treating contaminated soil. *Appl Environ Microbiol* **71**: 1202-1209.
- Singleton DR, Ramirez LG, Aitken MD (2009). Characterization of a polycyclic aromatic hydrocarbon degradation gene cluster in a phenanthrene-degrading *Acidovorax* strain. *Appl. Environ. Microbiol.* **75**: 2613-2620.
- Singleton DR, Sangaiah R, Gold A, Ball LM, Aitken MD (2006). Identification and quantification of uncultivated Proteobacteria associated with pyrene degradation in a bioreactor treating PAH-contaminated soil. *Environ Microbiol* **8**: 1736-1745.
- Singleton I, Merrington G, Colvan S, Delahunty JS (2003). The potential of soil protein-based methods to indicate metal contamination. *Appl. Soil Ecol.* **23**: 25-32.
- Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD *et al* (1985). Measurement of protein using bicinchoninic acid. *Anal. Biochem.* **150**: 76-85.
- Snijders APL, de Vos MGJ, Wright PC (2005). Novel approach for peptide quantitation and sequencing based on N-15 and C-13 metabolic labeling. *J. Proteome Res.* **4**: 578-585.
- Stingley RL, Khan AA, Cerniglia CE (2004). Molecular characterization of a phenanthrene degradation pathway in *Mycobacterium vanbaalenii* PYR-1. *Biochem Biophys Res Commun* **322**: 133-146.

- Suen WC, Haigler BE, Spain JC (1996). 2,4-Dinitrotoluene dioxygenase from *Burkholderia* sp. strain DNT: similarity to naphthalene dioxygenase. *J Bacteriol* **178**: 4926-34.
- Sueoka K, Satoh H, Onuki M, Mino T (2009). Microorganisms involved in anaerobic phenol degradation in the treatment of synthetic coke-oven wastewater detected by RNA stable-isotope probing. *FEMS Microbiol Lett* **291**: 169-74.
- Suflita JM, Davidova IA, Gieg LM, Duncan KE (2007). Anaerobic phenanthrene mineralization by a carboxylating sulfate-reducing bacterial enrichment. *ISME J.* **1**: 436-442.
- Sutherland JB (1992). Detoxification of polycyclic aromatic-hydrocarbons by fungi. *J. Ind. Microbiol.* **9**: 53-61.
- Takizawa N, Kaida N, Torigoe S, Moritani T, Sawada T, Satoh S *et al* (1994). Identification and characterization of genes encoding polycyclic aromatic hydrocarbon dioxygenase and polycyclic aromatic hydrocarbon dihydrodiol dehydrogenase in *Pseudomonas putida* OUS82. *J Bacteriol* **176**: 2444 - 2449.
- Tam NFY, Lei AP, Hu ZL, Wong YS (2007). Removal of fluoranthene and pyrene by different microalgal species. *Bioresour. Technol.* **98**: 273-280.
- Tittabutr P, Cho IK, Li QX (2011). Phn and Nag-like dioxygenases metabolize polycyclic aromatic hydrocarbons in *Burkholderia* sp. C3. *Biodegradation* **22**: 1119-1133.
- Torsvik V, Sorheim R, Goksoyr J (1996). Total bacterial diversity in soil and sediment communities - A review. *J. Ind. Microbiol.* **17**: 170-178.
- van der Meer JR, Devos WM, Harayama S, Zehnder AJB (1992). Molecular mechanisms of genetic adaptation to xenobiotic compounds. *Microbiol. Rev.* **56**: 677-694.
- van der Meer JR, Sentchilo V (2003). Genomic islands and the evolution of catabolic pathways in bacteria. *Curr. Opin. Biotechnol.* **14**: 248-254.
- van Elsas JD, Inceoglu O, Hoogwout EF, Hill P (2010). Effect of DNA extraction method on the apparent microbial diversity of soil. *Appl. Environ. Microbiol.* **76**: 3378-3382.
- Vinas M, Sabate J, Espuny MJ, Solanas AM (2005). Bacterial community dynamics and polycyclic aromatic hydrocarbon degradation during bioremediation of heavily creosote-contaminated soil. *Appl. Environ. Microbiol.* **71**: 7008-7018.
- von Bergen M, Benndorf D, Vogt C, Jehmlich N, Schmidt Y, Thomas H *et al* (2009). Improving protein extraction and separation methods for investigating the metaproteome of anaerobic benzene communities within sediments. *Biodegradation* **20**: 737-750.
- von Bergen M, Jehmlich N, Schmidt F, Taubert M, Seifert J, Bastida F *et al* (2010). Protein-based stable isotope probing. *Nat. Protoc.* **5**: 1957-1966.
- Walter U, Beyer M, Klein J, Rehm HJ (1991). Degradation of pyrene by *Rhodococcus* sp Uw1. *Appl. Microbiol. Biotechnol.* **34**: 671-676.

- Wang J, Van Nostrand JD, Wu LY, He ZL, Li GH, Zhou JZ (2011). Microarray-based evaluation of whole-community genome DNA amplification methods. *Appl. Environ. Microbiol.* **77**: 4241-4245.
- Warshawsky D, Cody T, Radike M, Reilman R, Schumann B, LaDow K *et al* (1995). Biotransformation of benzo[a]pyrene and other polycyclic aromatic hydrocarbons and heterocyclic analogs by several green algae and other algal species under gold and white light. *Chem Biol Interact* **97**: 131-48.
- Widdel F, Rabus R (2001). Anaerobic biodegradation of saturated and aromatic hydrocarbons. *Curr. Opin. Biotechnol.* **12**: 259-276.
- Willison JC (2004). Isolation and characterization of a novel sphingomonad capable of growth with chrysene as sole carbon and energy source. *FEMS Microbiol Lett* **241**: 143-50.
- Wilson MS, Bakermans C, Madsen EL (1999). In situ, real-time catabolic gene expression: Extraction and characterization of naphthalene dioxygenase mRNA transcripts from groundwater. *Appl Environ Microbiol* **65**: 80-87.
- Woese CR, Kandler O, Wheelis ML (1990). Towards a natural system of organisms - proposal for the domains archaea, bacteria, and eucarya. *Proc. Natl. Acad. Sci. USA.* **87**: 4576-4579.
- Wolfe MD, Lipscomb JD (2003). Hydrogen peroxide-coupled *cis*-diol formation catalyzed by naphthalene 1,2-dioxygenase. *J. Biol. Chem.* **278**: 829-835.
- Wolfe MD, Parales JV, Gibson DT, Lipscomb JD (2001). Single turnover chemistry and regulation of O-2 activation by the oxygenase component of naphthalene 1,2-dioxygenase. *J. Biol. Chem.* **276**: 1945-1953.
- Yamazoe A, Yagi O, Oyaizu H (2004). Degradation of polycyclic aromatic hydrocarbons by a newly isolated dibenzofuran-utilizing *Janibacter* sp strain YY-1. *Appl. Microbiol. Biotechnol.* **65**: 211-218.
- Yang YJ, Chen RF, Shiaris MP (1994). Metabolism of naphthalene, fluorene, and phenanthrene - preliminary characterization of a cloned gene-cluster from *Pseudomonas putida* NCIB-9816. *J Bacteriol* **176**: 2158-2164.
- Yanisch-Perron C, Vieira J, Messing J (1985). Improved M13 phage cloning vectors and host strains - nucleotide-sequences of the M13mp18 and pUC19 Vectors. *Gene* **33**: 103-119.
- Ye DY, Siddiqi MA, Maccubbin AE, Kumar S, Sikka HC (1996). Degradation of polynuclear aromatic hydrocarbons by *Sphingomonas paucimobilis*. *Environ. Sci. Technol.* **30**: 136-142.
- Yergeau E, Arbour M, Brousseau R, Juck D, Lawrence JR, Masson L *et al* (2009). Microarray and Real-Time PCR analyses of the responses of high-arctic soil bacteria to hydrocarbon pollution and bioremediation treatments. *Appl. Environ. Microbiol.* **75**: 6258-6267.

- Yu CL, Liu W, Ferraro DJ, Brown EN, Parales JV, Ramaswamy S *et al* (2007). Purification, characterization, and crystallization of the components of a biphenyl dioxygenase system from *Sphingobium yanoikuyae* B1. *J. Ind. Microbiol. Biotechnol.* **34**: 311-324.
- Zheng ZM, Obbard JP (2002). Polycyclic aromatic hydrocarbon removal from soil by surfactant solubilization and *Phanerochaete chrysosporium* oxidation. *J. Environ. Qual.* **31**: 1842-1847.
- Zhou HW, Guo CL, Wong YS, Tam NFY (2006). Genetic diversity of dioxygenase genes in polycyclic aromatic hydrocarbon-degrading bacteria isolated from mangrove sediments. *FEMS Microbiol Lett* **262**: 148-157.
- Zhou NY, Fuenmayor SL, Williams PA (2001). *nag* genes of *Ralstonia* (formerly *Pseudomonas*) sp. strain U2 encoding enzymes for gentisate catabolism. *J Bacteriol* **183**: 700-8.
- Zinger L, Coissac E, Choler P, Geremia RA (2009). Assessment of microbial communities by graph partitioning in a study of soil fungi in two alpine meadows. *Appl. Environ. Microbiol.* **75**: 5863-5870.
- Zinger L, Gury J, Giraud F, Krivobok S, Gielly L, Taberlet P *et al* (2007). Improvements of polymerase chain reaction and capillary electrophoresis single-strand conformation polymorphism methods in microbial ecology: Toward a high-throughput method for microbial diversity studies in soil. *Microb Ecol* **54**: 203-216.
- Zylstra GJ, Kim E, Goyal AK (1997). Comparative molecular analysis of genes for polycyclic aromatic hydrocarbon degradation. *Genet Eng (N Y)* **19**: 257-69.

Liste des annexes

Certaines annexes sont trop volumineuses pour être ajoutées ici.

Elles seront présentées dans le CD qui accompagne ce mémoire de thèse.

La liste des annexes disponibles est néanmoins détaillée ici.

Annexe 1 : Tableau des amorces et des programmes d'amplification (fichier Excel)

Annexe 2 : Scripts utilisés pour traiter les données SSCP dans Scilab et R (fichier Word)

Annexe 3 : Script utilisé pour traiter les données de séquençage dans mothur (fichier Word)

Annexe 4 : Données supplémentaires de qPCR : courbes d'inhibition, abondances relatives des phylums pour l'ADN ^{12}C et l'ADN ^{13}C (fichier Word)

Annexe 5 : Données de séquençage - numéro d'accession et affiliation des séquences (fichier Excel)

Annexe 6 : Attribution des peptides de l'analyse protéomique (fichier Excel)

Abstract

Polycyclic aromatic hydrocarbons (PAHs) are ubiquitous compounds produced by incomplete combustion of organic matter. They are a source of environmental pollution, especially associated to oil product exploitation, and represent a threat for living organisms including human beings because of their toxicity. Many bacteria capable of degrading PAHs have been isolated and studied. However, since less than 5% of soil bacteria can be cultivated in the laboratory, bacterial species able to degrade PAHs *in situ* have been poorly studied. The first goal of this study was to identify bacteria that degrade PAHs in soil using culture-independent molecular methods. To this end, a strategy known as stable isotope probing has been implemented based on the use of phenanthrene, a three rings PAH, in which the natural isotope of carbon was replaced by ^{13}C . This molecule has been introduced as a tracer in microcosms containing soil from a constructed wetlands collecting contaminated water from highway runoff. Bacteria having incorporated the ^{13}C were then identified by 16S rRNA gene sequence analysis after PCR amplification from labeled genomic DNA extracted from soil. The results show that so far little studied Betaproteobacteria, belonging to the genera *Acidovorax*, *Rhodoferrax*, *Hydrogenophaga* and *Thiobacillus*, as well as Rhodocyclaceae, were the key players in phenanthrene degradation. Predominance of Betaproteobacteria was established thanks to quantitative PCR measurements. A dynamic analysis of bacterial diversity also showed that the community structure of degraders depended on phenanthrene bioavailability. In addition, the phylogenetic diversity of ring-hydroxylating dioxygenases, enzymes involved in the first step of PAH degradation, has been explored. We detected new sequences, mostly related to dioxygenases from Sphingomonadales and Burkholderiales. For the first time, we were able to associate a catalytic activity for oxidation of PAHs to partial gene sequences amplified from soil DNA, by constructing hybrid enzymes and assaying their activity

The results obtained and the tools implemented in this study may be used to develop methods for the diagnostic and monitoring of pollutant biodegradation in processes such as bioremediation of PAHs contaminated sites.

Key words: bacterial diversity, isotopic labeling, 16S rRNA genes, ring-hydroxylating dioxygenases, phylogenetic and functional analysis, qPCR, proteomics, betaproteobacteria, PAHs, phenanthrene, biodegradation.

Résumé

Les hydrocarbures aromatiques polycycliques (HAP) sont des composés ubiquitaires issus de la combustion incomplète de matières organiques. Surtout liées à l'exploitation des produits pétroliers, ils sont à l'origine de pollutions de l'environnement, car ce sont des composés toxiques pour les êtres vivants et pour l'homme en particulier. De nombreuses bactéries capables de dégrader les HAP ont été isolées et étudiées, mais celles qui les dégradent *in situ* sont mal connues. En effet, moins de 5% des bactéries du sol sont cultivables en laboratoire. Le premier objectif de cette étude était d'identifier les bactéries qui dégradent les HAP dans le sol par des méthodes moléculaires indépendantes de la culture. A cette fin, une stratégie de marquage isotopique *in situ* a été mise en œuvre. Elle repose sur l'utilisation du phénanthrène, un HAP à trois cycles, dans lequel l'isotope naturel du carbone a été remplacé par le ^{13}C . Cette molécule a été introduite comme traceur dans des microcosmes contenant du sol provenant d'un bassin de rétention des eaux de ruissellement d'autoroute. Les bactéries capables de métaboliser le HAP ont incorporé le ^{13}C et ont ainsi été identifiées par séquençage des gènes d'ARNr 16S, amplifiés à partir de l'ADN marqué extrait du sol. Les résultats montrent que des Betaprotéobactéries peu étudiées à ce jour, appartenant aux genres *Acidovorax*, *Rhodoferax*, *Hydrogenophaga* et *Thiobacillus*, ainsi que des Rhodocyclaceae, étaient les principaux acteurs de la dégradation du phénanthrène. La prépondérance des Betaprotéobactéries a été établie par des mesures de PCR quantitative. Une analyse dynamique de la diversité bactérienne a montré que celle-ci changeait en fonction de la biodisponibilité du phénanthrène. En outre, la diversité d'arène-dioxygénases impliquées dans la dégradation des HAP a été explorée sur le plan phylogénétique et fonctionnel. Nous avons ainsi détecté des séquences nouvelles, pour la plupart apparentées à des dioxygénases de Sphingomonadales et de Burkholderiales. Grâce à la construction et à l'expression d'enzymes hybrides, il a été possible, pour la première fois, d'associer une activité catalytique d'oxydation des HAP à des séquences partielles de gènes, amplifiées à partir de l'ADN du sol. Les résultats obtenus et les outils mis au point dans cette étude pourront servir à développer des méthodes de diagnostic et de suivi de la biodégradation de polluants, par exemple dans le cadre d'opérations de bioremédiation de sites pollués par les HAP.

Mots clés : diversité bactérienne, marquage isotopique, gènes d'ARNr16S, arène-dioxygénases, analyse phylogénétique et fonctionnelle, qPCR, protéomique, betaprotéobactéries, HAPs, phénanthrène, biodégradation.